NASA

World Wide Web Best Practices

Draft 2.0

Prepared By:

NASA Web Best Practices Team
Chartered By the NASA Chief Information Officer and the Principal Center for Communications Architecture

Last Update: August 29, 2000

Table of Contents

Team Members

DJ Byrne — Jet Propulsion Laboratory

John Clack — Johnson Space Center

Jayne Dutra — Jet Propulsion Laboratory

Jay Dyson — Jet Propulsion Laboratory

Gaye Graves — Ames Research Center

Kelly Heidman — Glenn Research Center

Jeanne Holm — Jet Propulsion Laboratory

Nicole Masjedizadeh — Ames Research Center

Annette Randall — Ames Research Center

Peter Scott — Jet Propulsion Laboratory

Eric Tauer — Jet Propulsion Laboratory

Sue Wheaton — Dryden Flight Research Center

Executive Summary

The Internet and the World Wide Web (WWW) have become an integral component of the provision of NASA information to academia, industry, and the public, and as an enabling set of technologies for the internal NASA information technology environment. This document provides a set of web “Best Practices” (WBP) for all NASA entities engaged in the development and maintenance of World Wide Web resources. The Best Practices defined in this document are applicable to all NASA Enterprises, Centers, and project and program offices and have been created under the auspices of the NASA Chief Information Officer (CIO).

This Best Practices document is organized into the following chapters:

1. Site Design
2. Page Design
3. Site Accountability
4. Site Security
5. Accessibility
6. Web Marketing
7. Site Search
8. Authoring Tools

These Best Practices facilitate consistent, reliable, and efficient deployment of WWW services throughout the Agency. These criteria should be viewed as baseline recommendations. References to applicable government and NASA policies and standards are included to highlight “requirements.” This document will continue to be reviewed and revised by the NASA WWW community.

Introduction

Transmitting information to NASA partners, the academic community, industry and the public is a fundamental component of the NASA charter. The Internet and the World Wide Web (WWW) are integral elements in this task. These technologies also serve as a core piece of the internal information technology (IT) infrastructure on which NASA relies to carry out its mission. In this context, the set of technologies, content, processes, and policies that support the delivery of WWW-based services should be viewed as a significant part of the NASA environment and approached with the same commitment to excellence as all other NASA activities.

Purpose of Best Practices Guidelines

The Best Practices outlined in this document serve as a guideline to all NASA entities engaged in the development and maintenance of WWW resources. These guidelines facilitate consistent, reliable, and efficient deployment of WWW services throughout the Agency. These criteria should be viewed as baseline recommendations. References to applicable government and NASA policies and standards are included to emphasize "requirements." These best practices will continue to be reviewed and revised by the NASA WWW community.

The current document and associated information is available at:
http://nasa-wbp.larc.nasa.gov/

Scope

The guidelines defined in this document are applicable to all NASA Enterprises, Centers, and project and program offices. These guidelines are created and maintained under the auspices of the NASA Chief Information Officer (CIO).

Guideline Reviews

As the technology and government policy evolve, these guidelines will be reviewed by a NASA agency team on an annual basis. The review date is displayed on the title page of the document. Documents with review dates exceeding four months may contain obsolete or inaccurate information; the most recent version can be obtained from the NASA WWW Best Practice site at http://nasa-wbp.larc.nasa.gov/.

1.0 Site Design

This chapter describes high-level site design aspects, often referred to as *information architecture*. The two most important factors in site design are:

1. To identify the main goals and objectives of the site in order to effectively communicate its purpose.
2. To identify the target audience in order to understand the motivations, goals, behaviors, and technical requirements of that target audience.

To meet the intended audiences' goals, you must develop a specific purpose and articulate this to the intended audience, either explicitly or implicitly. By identifying primary goals and objectives, you can identify the site's content and its organization. A clear goal statement should be included in the content—possibly as an overview of the site's contents.

The site design should accommodate its goals and priorities. For example, does a site give priority to publishing recent developments, such as the NASA Home Page (http://www.nasa.gov) or to providing access to catalogs of information, such as the National Space Science Data Center (NSSDC) Web Site (http://nssdc.gsfc.nasa.gov/)? If the main purpose of the site is to publish recent developments, you need to allocate space within the design for new material and for easy updates. If the context does not require frequent updates, but encompasses extensive data holdings, then you need to focus on structure and network capacity.

Determine How Your Audience Will Use the Site

To communicate effectively with the intended audience, you must understand their motivations for visiting a site and their behavior in navigating and using the information. A clear understanding of how the target audience will use a site helps you determine what content, level of detail, format, and navigational aids you should employ.

An intranet site, for example, accommodates an internal audience by providing well-organized content that is quickly accessed. A public site, on the other hand, must draw users into the site by stating the site's purpose and preview of the site's contents.

Many studies have been done to determine NASA's target audience. As a result, we know that the Agency has a very broad, but also segmented, audience divided into five groups:

1. **General Public and mass media:** These audiences typically have some general knowledge of NASA's missions and accomplishments.

2. **Research communities beyond NASA: industry, scientific and technical press:** These audiences typically possess in-depth knowledge about specific subjects.
3. **Government groups such as Congress, the Presidential Administration, and other agencies:** These audiences have varying levels of familiarity with NASA subject matter.
4. **Internal:** Primary focus may be on a Center, a project, or small organizational group.
5. **Academia:** Teachers, administrators and students. Recently this group has become a much higher priority for both the Federal Government Administration and NASA. NASA employees who are primarily interested in reaching these groups must bear in mind that educational material needs to be tied to public education curricula. Employees should work with their Centers' educational outreach groups to ensure that web sites meet these needs.

References:

- Yale Style Manual
  http://info.med.yale.edu/caim/manual/sites/site_elements.html

- WWW Federal Consortium Guidelines and Best Practices
  http://hoohoo.ncsa.uiuc.edu:7777/consortium_info/agency_policy_guidelines/agency_policy.html

1.1 Structure and Site Organization

1.1.1 Information Architecture

Information architecture refers to how information is structured to accommodate the audience and the site's purpose. Information architecture is determined by how users will browse and navigate the site.

Information architecture determines the site's usability. The features you must consider include the following:

- **Browse**

How the user browses your site determines the user's "experience" of the site. Some sites provide little distinct organizational flow, but rather offer materials in a "find and seek" approach. Browsing is facilitated by providing several methods of accessing the information: organizing content in alphabetical order, by keyword categories and/or table of contents. Letting the user "discover" the site's elements is well suited to children's materials on an outreach site, where the user might be exploring and discovering concepts.

- **Classification**

Some sites are based on a standard taxonomy (or classification). For example, a NASA site might use the NASA Thesaurus for its taxonomy. Large sites with diverse content can be structured by categories, providing various subject headings to encompass the variety of topics. A well-known example of this type of structure is the Yahoo site http://www.yahoo.com.

- **Cluster**

Clustering different aspects of a central idea or element is useful for a small site with a very specific focus (for example, car buying: costs, locations, financing, service).

- **Flowchart**

To show steps in a sequential order or a process use a flow diagram; for example, show the steps in how to paint a car: sanding> primer coat> final top coat.

- **Hierarchical**

Hierarchical order presents information in main categories and subcategories that are structured vertically. This organizational pattern is used for sites with dense information content.

- **Navigation Pointers**

Navigation pointers like labeled buttons, descriptive links, highlighted terms help users navigate the site. Present clusters of information with clear pointers either as text or as graphic pointers to lead the user to each area. The information is logically grouped into chunks and identified with descriptive labels and icons.

1.1.2 Guidelines for Various Structural Formats

The two most common structural formats are hierarchical structure and classification structure. Guidelines for using each follow:

Hierarchical Structures (also known as tree structures)

- For practicality and efficiency, analyze the functional and aesthetic aspects of the site.
- For a directory structure provide a balance of topics; avoid too few choices or too many choices at each menu level. The number of primary levels can be concise and relevant by presenting original versus re-purposed content.
- The hierarchy should represent some logical division of the overall content, keeping important information toward the top. Establish a hierarchy of importance and divide pages/information/content into logical units. Use the hierarchy to show relationships among topics.

- To determine the effectiveness of your organization, use a metric that counts the number of mouse clicks necessary to locate information. Most experts suggest the depth should be between one and three clicks.

Classification Structures

- Use a consistent vocabulary that reflects the classification structure.
- Develop a structure based on a well-established, standard classification scheme; example schemes include the
  - NASA Thesaurus
  - Engineering Information (EI) Classification Codes
  - Mathematics Subject Classification
  - Universal Decimal Classification (UDC)
- Plan for additional information clusters as the site grows. Review the structure of your subject headings and subheadings to ensure future site "extensibility." By planning a site for growth, users become familiar with your organizational structure and therefore are not confused when the site grows.
- Avoid general or heterogeneous subject areas. Include descriptive subheadings.

References:

- "Net Profit in a Post Modem World," by Jamie McKenzie
  http://www.fromnowon.org
- Jakob Nielsen's Alertbox, Oct. 17, 1999: "Prioritize: Good Content Bubbles to the Top"
  http://www.useit.com/alertbox/991017.html
- Web Reviews' Web Architect, "Label Laws by Lou Rosenfeld"
  http://webreview.com/pub/96/03/29/webarch/index.html
- San Diego City Schools' Patterns Project, "A Taxonomy of Information Patterns"
  http://projects.edtech.sandi.net/staffdev/tpss98/patterns-taxonomy.html
- Support Initiative for Multimedia Applications (SIMA). Sue Cunningham at Manchester Visualization Centre
  mailto:sue.cunninghame@mcc.ac.uk
- Yale Style Manual
  http://info.med.yale.edu/caim/manual/sites/site_design.html

- NCSA TRG Review of Web Styleguides
  http://www.ncsa.uiuc.edu/edu/trg/styleguide/index-r.html#3

- NASA Thesaurus
  http://www.sti.nasa.gov/thes1.htm

1.2 Site Navigation

A well-designed site simplifies the task of locating information and facilitates the user's ability to move from one page to another. The architecture of the site provides the users a means of identifying their current location in the site's space and provides navigation tools to guide them from one location to another.

To assist users in navigating through a site's content, follow these guidelines:

Navigation should be easily learned and should remain consistent throughout the site.

- Make your navigation scheme transparent to your users. Users will not spend the time to learn a complicated scheme. Remember you have only one chance to make a good first impression.

- Be consistent in the placement and design of navigation elements. Users expect that navigation buttons and bars be displayed in the same place on every page. Consistency builds the users' trust and enhances the quality of the experience.

Provide visual context

- Visual techniques like templates and style sheets, and items such as last updated or modified date stamps, provide important clues so users know where they are in the site. Include a "Home" button on each page, giving users quick access the site's top-level information.

Minimize reliance on browser navigational buttons

- Create site-specific navigational links on each web page so users can intuit where they are in the site regardless of their entry point.

- Do not confuse users by changing expected browser behaviors. For instance, if they expect the browser's back button to take them to the previous page, don't break this rule without a good reason.

Provide clear visual messages and labels

- Interface design is about visual guidance. How navigation options are presented is closely tied to how usable they are. If these options are hidden, difficult to find, or look too much like text or are otherwise visually confusing, your users will have trouble navigating the site.
- When selecting labels, use terminology familiar to your users. Insider jargon is a barrier to clear communication. Good labeling is based on common sense and user sensitivity.

Provide a search mechanism

- Sites should provide some type of search interface or site index to help users quickly access information that may be located deep within the site. More information on local site search is available in Chapter 7, Site Search.

References:

- NCSA TRG Review of Web Styleguides
  http://www.ncsa.uiuc.edu/edu/trg/styleguide/index-r.html#6
- Yale Style Manual
  http://info.med.yale.edu/caim/manual/intro/purpose.html
- Sun Styleguide
  http://www.sun.com/styleguide/others/Navigation.html
- *Web Navigation: Designing the User Experience*, Jennifer Fleming

1.3 File Formats and Data Coding

1.3.1 File Format Types

The file formats used on a web site have to satisfy the often-conflicting demands of design, function, and user needs. The primary criterion for selecting file formats should be whether the target audience can easily use a specific file format. For publicly accessible sites, select file formats that are widely available and commonly supported by most popular web browsers across all major platforms. Fortunately, the number of file formats directly handled by most web browsers is constantly growing.

Interchangeable and downloadable file formats

- Provide a choice of formats so users can find a file format compatible with their software. Adobe Portable Document Format (PDF), Rich Text Format (RTF), Microsoft Word, Excel, or GIF and JPEG are commonly used formats. In closed

communities, such as intranets, you have greater control over the applications available at the desktop, and additional file formats may be supported, e.g., Informed Manager or Microsoft Project. However, be aware of the Center's baseline software suite, and avoid formats that are not supported by the Center's information technology group. Also be aware that delivering the same information in multiple formats can increase maintenance and increase the possibility of information being out of sync.

Formats requiring special viewers (or "plug-ins")

- Plug-ins, like a player or special viewer, are required to ensure some files display properly. For example, files saved in the QuickTime, RealAudio, or Shockwave Flash formats can be viewed on players that are shipped and installed with the latest versions of Netscape and Internet Explorer for both the Mac and PC platforms. For the widest public access to your site, select file formats that are widely supported by the popular web browsers.

- Your web pages should be viewable even if the plug-in is not installed. A good practice is to provide a link to the plug-in download page. For information critical to understanding or navigating the site, provide the same information in an alternative format, such as text. Users become frustrated when a NASA site requires the installation of plug-ins or other software that the users need to install before they can access information.

1.3.2 Coding and Programming Languages

HTML

Content on the web is manipulated by computer codes ranging from the very simple such as HTML to the complex such as Java. Different browsers and different versions of the same browser may not support different forms of content and language in a consistent manner. For example, earlier versions of HTML did a poor job of implementing frames; however, now that frames are a part of the HTML 4.0 standard, any 4.0-compliant browser must implement them, so cross-browser compatibility will be less of an issue.

Cascading Style Sheets

Cascading Style Sheets (CSS) are another example of a feature that is not yet mature. CSS standards are changing and their various style properties never seem to work the same on all browsers. Nevertheless, as the Internet moves away from HTML toward XHTML, tomorrow's web developer will need to learn how to use CSS, since today's formatting tags won't work with XHTML.

JavaScript

JavaScript is not considered a programming language but one form of code that is widely used on the web. Most browsers support JavaScript 1.l, and if the code is written well, the script won't fail when viewed on most browsers. The more complex the language, the more important that you know how to write good code. Equally important is testing the compatibility of your code on different platforms and browsers.

You can develop complex content through simple, robust HTML coding that is read by various browser versions and platforms. However complex coding depends on your mission, your skills, your audience, and your free time.

References:

- World Wide Web Consortium
  http://www.w3.org/
- MSDN Web Workshop - HTML for Beginners
  http://msdn.microsoft.com/workshop/author/html/beghtml.asp
- A master style sheet compatibility chart
  http://webreview.com/wr/pub/guides/style/mastergrid.html
- JavaScript compatibility guide and chart
  http://webreview.com/pub/1999/10/29/feature/index3b.html
- Any good reference or technical book, such as the O'Reilly series.

1.4 Maintainability

1.4.1 The Art of Reuse

Content must be developed for reuse to reduce total site maintenance. This can be done in several ways:

- Use short, readable HTML documents; this will aid in locating problems and making corrections as the content evolves.
- Format your HTML code with indents and blank lines separating sections so that it is easy to read. Be consistent in your formatting so that changes can be incorporated quickly and easily updated by someone. The comment tag is also useful when labeling sections of an HTML file.
- Provide a diagram of site architecture and keep it updated. This diagram helps you remember where documents are stored if you have not worked on an individual site in some time. A diagram is a useful tool to pass on to

another person, who may be assigned to maintain the site.

- Store frequently used content components like images, headers, footers, etc. in shared, centralized locations, rather than replicating the same content in multiple locations on the server. This practice reduces duplicate maintenance for similar components and conserves disk space as well as reduces download time.
- Date your pages individually, and edit the date every time a change is made.
- Use global templates/style sheets or server-side includes to reduce the number of files that need to be maintained.
- Avoid using browser-dependent features that require two or more versions of HTML code.

1.4.2 Links

Links within a web site are probably the single largest maintenance issue and often the greatest source of frustration for users. To minimize invalid or broken links, follow these practices:

- Strike a balance between the number of links offered and the quality of those links. The greater the number of links, the greater the probability that someone will move a target document and break a link.
- Do not move documents or change uniform resource locators (URLs) too often. Frequent changes increase the possibility of broken links or breaking bookmarks that visitors have made to a site.
- If the site must be moved then post a pointer page at the old URL, or use server-side redirects to send the user to the new site. A pointer page notifies the user that the site has moved to a new URL, gives the location of the URL, and provides a direct link to the new page location. Generally pointer pages remain at the old site for approximately 3 to 6 months. The second option, server-side redirects, is configured within the web server and redirects the user to the new site while bypassing the old site altogether. This alleviates a stop at the old site, as required by a pointer page. However, sometimes users do not realize they have been redirected to another URL, and they will not bookmark or modify their site link to the new location.
- Use relative URLs links instead of absolute links between pages within a site will simplify the administrative aspect of relocating the content,

should that be necessary.

- Use some form of automation, such as Momspider or Linkbot, whenever possible to identify broken links that will inevitably occur. Some text editors include link validation in their tool suite. Plan site checks on a regular basis that are part of a maintenance schedule.

1.4.3 Automation

Even when care has been taken to minimize site maintenance manual checks may not be enough to ensure proper care of the site. The best way to effectively maintain a site is to use automation to assist in identifying problems and in creating content.

- **Link validation**: Rather than attempting to manually scan hundreds and even thousands of documents for broken links, use a link validation tool, and run these tools at regular intervals to identify and report broken links.
- **HTML validation**: Use these tools to identify HTML coding errors or HTML features that may not be supported by all browsers. These tools should be used before site release and periodically according to your maintenance schedule.
- When practical, use automation to generate content based on selected inputs, thereby minimizing the potential for human error. Obviously, sites that generate their content automatically from databases or other data sources eliminate or significantly reduce the amount of page and site maintenance; only the information stored in the database needs to be maintained. Keep in mind, however, that to have a database site set up can be expensive and is usually not cost effective for small sites or for sites that do not require frequent changes.

1.4.4 Feedback

Despite all attempts to ensure that links are accurate, errors will occur. Make sure that users can send feedback to site personnel so these problems can be corrected.

1.4.5 Variable Data

Users return to a site looking for fresh content. To generate return visits, change your site as often as possible within the boundaries of your time and budget. However, make changes that won't confuse your users. it easy for users to identify changes.

For sites that contain static information, avoid using data that will soon be outdated. For example, rather than state "the software was developed five years ago," state "the software was developed in 1995."

References:

- The Sev Guide to Web Design: Maintenance
  http://www.sev.com.au/webzone/design/maintenance.asp
- NCSA - TRG: Review of Web Styleguides
  http://www.ncsa.uiuc.edu/edu/trg/styleguide/index-r.html#9
- What Makes a Great Web Site?
  http://www.webreference.com/greatsite.html
- Art and the Zen of Web Sites
  http://www.tlc-systems.com/webtips.shtml
- See the "Guidance on Implementation of NASA Web Site Privacy Statement" in Appendix B.

1.5 NASA Integrated Information Technology Architecture

- The NASA Integrated Information Technology Architecture (NIITA) document is an evolving document that establishes a set of goals and guidelines that provide a blueprint for NASA IT service providers and, as such, defines a common, vendor-independent framework for design, integration and implementation of IT systems.

  In this document it states that,

  "NASA has … embraced the use of the World-Wide-Web (WWW) Browser as a 'universal' client and can be considered a leader in the use of WWW clients and servers for the dissemination and retrieval of information."

  The NIITA document also states that,

  "NASA has thoughtfully considered the idea of an 'intranet' and what it means for the Agency and the industry-at-large. The general consensus within the Agency is that an intranet is the structured use of the open, scalable technologies of the Internet (such as those listed in the preceding paragraph) to do the work of an Enterprise in a secure and reliable manner."

- Although the NIITA does not provide specific guidance with respect to web site design, it does define a set of basic principles that would generally apply:
  - All IT system design must be customer driven.
  - IT solutions should demonstrate a commitment to a standards-based, modular design.
  - IT solutions should utilize stable, commercially available solutions if at all possible.
  - IT solutions should leverage existing IT investment.
  - Design objectives should give preference to general-use solutions.
  - Design objectives should give preference to a simpler IT infrastructure that anticipates multiple, and emerging, end-user devices.

References:

- NASA-STD-2814
  http://www.hq.nasa.gov/office/cio/standards/2814.doc

2.0 Page Design

While site design focuses on issues that affect an entire site, web page design focuses on issues specific to individual pages or sections of pages. Web page design incorporates all areas of web development, ranging from organizing the flow of content for the user to creating the graphic design, multimedia and programming—all enhancing the user's experience.

2.1 Logical vs. Physical Markup

A well-designed page helps users locate information and facilitates the users' ability to move from one page to another. While page layout is a key factor in accomplishing this goal, simple and intuitive navigation plays a key role in enhancing usability.

Use logical markup in favor of physical markup.

- HTML includes tags for both *logical* and *physical* markup. Logical markup specifies the nature of a piece of text: such as a heading, a paragraph, an important word, a book title. Physical markup specifies how the text will look on screen: large font, bold face, italics.

  Whenever possible, use logical markup instead of physical markup to let the browser choose how to display the specified text. Sometimes the desired appearance (e.g., boldface) may not be available in a particular browser (e.g., on a character terminal or text-to-speech reader). In addition, the logical elements are generally more meaningful for parsing of the document. Declaring that something is a level 1 heading will be more useful than saying it is in bold, courier, and a 24-pt. font.

  The number of logical markup tags is presently limited, but XML (eXtensible Markup Language) will allow authors to invent their own tags and, in combination with style sheets, specify their preferred appearance.

  A good introduction to XML is available at:
  http://msdn.microsoft.com/xml/general/intro.asp

  Design tip: You may someday need to dissect the structure of all your pages by machine rather than by a browser. This activity will be easier if the structure is logical and clear.

2.2 Create a Consistent Look

Maintain a consistent appearance and navigation style throughout the web site. Consistency helps users identify your site, realize when they've left it, and navigate it

successfully. Elements that should be consistent on each page include the overall structure, colors, backgrounds, headers, footers, sidebars, and navigation cues.

Maintaining consistency is easier with careful planning, server-side includes, templates, scripts, and style sheets, etc., which will be discussed later in this section.

Design tip: Although from a hierarchical perspective, you may view your site as starting from the home page, many users will enter your site through a link to an intermediate page. Users should be able to immediately identify your site and recognize subsequent pages that are part of it.

2.2.1 Structure

Use logical divisions consistently.

- Each document comprises headings, text paragraphs, lists, tables, and other logical divisions. Using these elements consistently on all pages is the first step toward a uniform look.

2.2.2 Colors

Use black text on default or solid white background.

- Each browser has a default set of colors that are used for the page backgrounds and text. Typically the page has a gray or white background, black text, blue hypertext links that have not been visited, and purple for visited links. Pages with white backgrounds have been found to be the easiest to read. White background color in the HTML tags displays best on the greatest number of browsers.

  Some browsers allow the page author to modify default colors; it is tempting to give a document or site a unique look by doing so. However, finding color combinations that are readable on all monitors can be difficult. Changing the visited/unvisited colors may confuse viewers.

  If you do change the background color, specify text colors that do not conflict with the reader's default Color choice is an important consideration when attempting to create a site that will be accessible to handicapped or color blind users. See Chapter 5, Accessibility Guidelines, for more detail.

  Background color may also affect the printability of your web pages. On sites that have dark backgrounds, offer a "printer friendly" version of pages, which uses a white background and default link colors.

2.2.3 Backgrounds

Avoid background images.

- Some browsers allow the page author to specify an image as a background for a page. Typically the image is tiled to fill the window area. However, background images should be avoided because they increase download time. Also, text can become unreadable if it overlays parts of the image. Consistent alignment between background images and the text or other elements is rarely possible on different platforms or browsers. The page can look distorted when viewed in a very large or very small window.

  If a background is used, the image size should be as small as possible, and the image subtle and low contrast. A possible exception is when using a margin along one edge of the page, perhaps in conjunction with a sidebar. However, in this case, if the window size is changed, the alignment of the sidebar text may be destroyed. For full screen windows, the margin may repeat in the middle of the page.

2.2.4 Headers

Page headers should be narrow and fast loading.

- Page headers are text and/or graphics that appear at the top of every page. (Page headers should not be confused with HTML heading tags.) Using a similar-looking header creates a consistent look and feel for your pages. Examples of headers might be a line of text, a logo, and a long narrow banner graphic, image map, or a navigation bar.

  The header can also act as a navigational element such as an image map, or a navigation bar. Navigation bars are often “sliced” into several clickable navigational images or buttons. They can even incorporate a path summary or “bread crumb trail” for site navigation, as at http://www.useit.com. Because it takes up valuable screen space, the header should only be a few pixels tall. A single line of text wrapped in a table with a color fill is a quick loading alternative to a graphic header. Any image that is used should be designed to load quickly, as it is one of the first items loaded in the document.

2.2.5 Footers

Use the footer for navigation elements and for site-wide information that is too long or repetitive to fit in the header.

- A footer is text or graphics that appear at the bottom of the page. As with the headers, footers should be similar throughout the web site. Footers might include site or agency logos, contact addresses, and feedback requests.

  Footers may also include Agency- or Center-required elements such as links to a privacy statement or a disclaimer, and listings of the site curator, NASA Responsible Official and date last updated.

2.2.6 Sidebars

Use sidebars judiciously, using headers and footers whenever possible.

- An alternative to both headers and footers is a sidebar, which is a vertical column of text located typically along the left edge of the page. Sidebars often contain the elements commonly found in headers and footers. The advantage of a sidebar is that, unlike a header, it does not push the main contents farther down the page and, unlike footers, it is visible when the page is first loaded.

  A sidebar allows prominent display of a large number of links or small, featured bits of information, without pushing other content completely off the first screen. The sidebar, therefore, can function as a navigational element. The disadvantage of using a sidebar as a navigation aid is that it reduces the width available for each line of text, which is particularly a problem on small monitors, laptops or for users who choose not to use their browser full screen. Also, text-only or speech-enabled browsers tend to display the sidebar before the content, particularly if the sidebar is in the left margin.

  It is worth noting that if tables are used to create a sidebar, the page may not load as quickly because most browsers do not render elements until the entire table has arrived. This problem can be addressed by making the sidebar itself a table while putting the rest of the content in a normal page body. If frames are used for the sidebar, consider the WBP recommendations that frames should by used judiciously.

2.2.7 Frames

In the past, frames have been used as a method to lay out pages, so repeated elements are available on every page. For example, a sidebar in a frame on each page could contain navigational elements. However, WBP recommends avoiding frames altogether.

Avoid using frames.

- Frames make it difficult for the user to bookmark a specific page. Frames also make the process of printing the page confusing for the novice user. Coding the HTML to drop the correct page into the appropriate frame may become muddled, making it more difficult for search engines to find the pages. And JavaScript rollover menus may not function correctly as the user's cursor travels from frame to frame. Most important frames do not comply with the disabilities act. These issues are a few of the reasons why we do not recommend the use of frames.

2.2.8 Navigating the Site

Site navigation exists on three levels: (1) site level, (2) section level, and (3) page level. Site level navigation focuses on the navigational elements that should appear on every page in the site: Help, Contacts, Search, and Home.

Section level navigation elements only appear on pages within a section. They link to information that is pertinent to the topics referenced in the section. These section elements might be incorporated by adding a second row of buttons below the site level navigation. Section level navigation can be further differentiated by the use of colors that are different than those used for the more primary site level navigation.

Page level navigation includes elements used to navigate within the page such as links that allow the user to jump down the page to a specific topic or location.

2.2.8.1 Using Navigational Elements

Be consistent in your use of navigational elements.

- To maintain a unified look for a site and to avoid confusion for the user, choose a consistent set of navigational elements and use them on every page. Navigational elements can be text or image links, a row of icons in the footer, a menu in a sidebar, a navigation bar at the top of the page or a path summary in the header. Regardless of the type of links used, every page should have a link to the "home page" or to a higher-level page in the same section. Also Center or Agency rules may require including links to Center,

Agency or other locations in the Agency hierarchy. Regardless of the type of navigational elements, it is crucial that they be placed in consistent locations on the pages.

Design tip: A user who arrives at your site from a search engine may start at some obscure page in the middle of the site and will need links back up to the major sections.

Design tip: Links should be given descriptive names. Do not simply create a link that reads that reads "click here." Links should be self-explanatory where possible and should be part of the normal sentence structure of the site. Descriptive links also increase the accessibility of a site for screen readers used by visually impaired users. And always remember to create an ALT tag for any buttons or clickable images.

2.2.8.2 Browser compatibility issues for navigational elements

As HTML matures and new versions of browsers begin to support more capabilities, it is tempting to create navigational elements that incorporate these new features. New techniques, such as JavaScript rollover buttons are easy to create using HTML editors and are becoming commonplace. Java Applet-based interfaces and DHTML rollover menus are also becoming commonly used.

While these new techniques may produce intriguing interfaces, they are not always compatible with older versions of browsers. Making content accessible to the user should be of a higher priority than creating intriguing interfaces. If the user has turned Java or JavaScript off in the browser's preferences, navigational elements based on these capabilities may not function. A good practice is to test all interfaces on different platforms (Mac, PC, and UNIX), as well as using several versions of both Netscape and Internet Explorer browsers.

See Section 2.7 on Cross-Browser Compatibility for more details.

2.3 Maintain a Consistent Look

Once you have designed a look for your site, you must maintain this design as pages are added or modified. This task is simplified by planning the site's design and using server-side includes, templates, scripts, and style sheets.

2.3.1 Planning

When creating or revising a site, spend the time planning before starting to write HTML. Consider the major topics and subtopics, the structure and the layout of a typical page, and the navigation among pages. Begin writing content for your web pages and start writing descriptions and keywords to include in your META tags. Well-chosen keywords and descriptions improve your ratings with search engines. See Section 2.3.2 on writing for the web for more suggestions.

If possible, get *written* sign-off on content and design from site stakeholders or customers. Settling such issues in advance will avoid the need to redo pages later. Develop a hierarchical chart of the site and submit this to your customers.

2.3.2 Writing for the Web

Most people find reading from the screen an uncomfortable experience. Scrolling down lengthy pages becomes tedious. As a result, most users tend to skim or "surf" the page, reading titles, hypertext links, bulleted lists, and short paragraphs. As an alternative, people may print out pages to read later off-line. For more on "How Users Read on the Web," visit Jakob Nielsen's commentary at: http://www.useit.com/alertbox/9710a.html

To write content that will be read by the impatient web surfer, it's best to:

- Include your most important information in the lead sentence and use short two or three sentence paragraphs.
- Keep page content brief, no longer than one-and-a-half screens.
- Use hypertext links to highlight important information and lead the reader through your content.
- Use well-placed links. Because people may enter a web site on any page, your information should not be dependent on concepts presented on "higher level" pages. However, using well-planned links may avoid the necessity of presenting redundant information.
- Keep titles for sections of home pages short, no more than one or two words, especially if they are used as button labels for navigating the web site.
- Use descriptive keywords for META tags.
- Always spell check content.

The Sun Microsystems' *Guide to Web Style (*http://www.sun.com/styleguide/) is an excellent reference for writing web content.

2.3.3 Server-Side Includes

Server-side includes (SSIs) are page features found on some servers that allow insertion of one document into another before the server ships the page over the network. For example, a footer that is the same on every page can be stored in a separate file and inserted by the server. When you modify the footer document, all pages that refer to it are automatically updated. SSIs, if available on the server, are an easy way to help maintain consistency. One disadvantage of SSIs is response time is slightly slower because it takes time to parse a document and insert the requested file(s). Also, some servers do not set a last-modified time (for the benefit of caches) on server-parsed documents.

2.3.4 Templates

Templates are valuable tools for creating consistency. They can be provided to page developers as a style guide for page layout. Comments in the template recommend font sizes, font types, and locations for navigational elements as well as for adding new content.

A template can also be a general example of the SSI technique. If a site has a large number of similar pages, each with some static and some variable content, a template can be used to describe the static portions of the page. Tokens placed within the template are examined by the server and replaced by variable contents as needed. For example, a page showing information about a person might have static headers and navigation icons and variable tokens representing name, address, and phone number. The server might fetch those variables from a database and insert them before shipping the page.

Using templates may require either custom programming or a third-party product.

2.3.5 Scripts

Another way to maintain page uniformity is to generate pages dynamically using scripts. Scripts allow you to generate a collection of pages using a set of arguments. For instance, an image catalog can be scripted so each page is uniquely identified by the image it displays then a server-side script determines what page the user wants and then creates it on demand. Every page looks similar because the same script makes each page. A site-wide change can be implemented simply by making changes to that script.

Using scripts typically requires custom programming and places extra load on the server. However, scripts efficiently handle infinite permutations of content extracted from a database.

2.3.6 Style Sheets

Style sheets are becoming the preferred way to describe and maintain page appearance. A style sheet specifies how different elements look on the page. A style sheet can be coded within the HTML page, or it can be a separate file linked to the page. A large site with many pages should have a separate file. Using a separate file allows you to change the overall style of a web site by only changing one file, rather than changing a number of individual HTML pages. A style sheet ensures that every time a particular element (say, a level 1 heading) is used, it always appears consistent.

Unfortunately, the current browsers do not fully support style sheets and are not consistent in their implementation. Therefore, you must research the elements of style intended for a site and make sure that these elements are compatible on all browser versions and platforms. Good places to research this are (http://www.webreview.com/guides/style/), (http://www.w3.org/) or (http://msdn.microsoft.com/workshop/default.asp).

Style sheets should improve with time, and the result will be a powerful tool.

2.4 Graphics

Graphics are used in web pages for navigation or as design elements—for logos, illustrations, and decoration. When using graphics, follow these guidelines:

- Keep file sizes small. Commercial software is available to “optimize” or reduce file sizes.
- Provide useful ALT attributes. ALT is an attribute in an HTML image tag that provides a meaningful text alternative when images are not loaded.
- Provide an alternate means of navigation that does not require images
- Remember that most search engines cannot read or index images.
- Never use images from other sites without permission.
- Specify WIDTH and HEIGHT attributes so the browser can render the document quickly.
- Consult and, if appropriate, link to NASA’s “Reproduction Guidelines for Use of NASA Images and Emblems”: (http://www.NASA.gov/gallery/photo/guideline.html)

2.4.1 Graphics File Formats

Several types of graphic file formats are used for the web. Choose a file format based on a specific use. JPEG (Joint Photographic Experts Group) files work well for photographic images. GIF (Graphic Interchange Format) files work well with line drawings and text images. Older browsers may not support other file formats, such as PNG. When in doubt, test the images on various browsers and platforms.

2.4.2 Icons

Icons used as navigational elements give a site a unique and consistent look. Icons can be used to illustrate as well as link to the major sections of the site. To avoid confusion, use the same icon for the same function.

It is difficult to create icons whose meaning is obvious to all users. A small picture cannot always convey an abstract topic (e.g., consider inventing an icon for the topic "alkalinity"). However, libraries of icons exist that can be downloaded and used. These are simple to acquire, but may not make a site unique. Consider hiring an artist to design appropriate icons.

Design tip: Always make icons small (both in pixel size and bytes) so they load quickly and do not overwhelm the page content.

2.4.3 Logos

A discreet logo on each page helps unify a site.

- An Agency, Center, or project logo is a good way to identify a site and its sponsors. The logo should load quickly and unobtrusively, as visitors are more interested in content than in a logo. Choose an alternative text (ALT tag) that describes or replaces the logo (i.e., "Laboratory for Atmospheres" rather than just "logo").

  Institutional requirements dictate how the logo of the Agency or Center should be displayed; consult the appropriate Public Affairs Office for correct usage of the logo or visit the "Reproduction Guidelines for Use of NASA Images and Emblems" web page: http://www.nasa.gov/gallery/photo/guideline.html.

2.4.4 Illustrations

Choose and produce illustrations with care.

- Images and graphs are essential for illustrating conceptual ideas and research data. However do not overload a web page with large color graphics. Carefully select images for their aesthetics and use them sparingly to create a clean, uncluttered look.

Keep image file size as small as practical by using these techniques:

- crop or shrink the image
- reduce the number of colors
- consider using Joint Photographic Experts Group (JPEG) instead of Graphics Interchange Format (GIF) for continuous tone imagery.
- use a thumbnail images as a link to the full-size image when displaying scientific data; the document will load more quickly and readers can select only the images they care about
- provide alternative text that describes the image and allows the reader to choose whether to load it.

2.4.5 Non-Functional Images

Avoid decorative images.

· Carefully selected images can be used as logos or banners to help create a unified look for a site. However minimize the number of purely decorative images, which add to page download time, and avoid the gratuitous use of images that serve no purpose at all.

2.4.6 Image Layout

HTML provides limited capabilities for page layout, when compared to desktop publishing software or word processors. This is deliberate; HTML is designed to be viewed on a variety of devices, ranging from the tiny monochrome screen of a palmtop to a high-resolution true-color monitor. HTML may also be "viewed" by a text-to-speech reader or by an autonomous search engine. Attempting to control every aspect of the on-screen presentation or layout may interfere with device and platform compatibility. See the Accessibility Guidelines Chapter for more information on improving device compatibility.

Do not create a page using one large image.

· Avoid creating an entire page as an image. It may look just right on the screen, but the slow download time will deter most visitors. In addition, an image cannot be indexed by most search engines and therefore will not be searchable.

To create the illusion of a single graphic, divide the image into several small graphics that are held in place by tables. Smaller graphics can be coded easily as image links and appropriate "ALT tags" can be added. Many

current HTML editors and image processors offer this image slicing capability. Keep in mind that although the screen will load quicker with several small images than with a single image, it does take time to download several images.

Be cautious with invisible spacers and other layout tricks.

- Some designers use invisible images (transparent images in GIF89a format, or images the same color as the background) to place page elements into the desired location. This trick is harmless if done properly. It is crucial to use only small images, supply alternative text attributes (ALT="spacer"), and view the result in a variety of browsers, with image loading disabled, and at several screen sizes. The presentation may differ on various browsers or platforms and may be not worth the effort required to produce it.

  As previously mentioned, you can create sophisticated layouts by dividing a larger image and putting the pieces into separate cells of a table. When doing this, follow the same rules as mentioned in the discussion about invisible spacer GIFs.

Design tip: Focus on the information conveyed by your page rather than its appearance. Useful content is timeless; nice appearance depends on current fashion.

2.5 Image Maps

Use image maps when necessary, and provide an alternate means of navigation.

- Image maps allow different parts of a single image to be linked to different documents on a server. Image maps are useful, even essential, in many applications. For example, clicking on a geographic map to select a country is a reasonable user interface style. An image map might also be used for a row of navigation icons.

  Image maps, however, can easily be overused, such as when the image is simply a picture of text and various phrases are links to different documents. The same result can be achieved using HTML, unless a specific style of text is desired to maintain consistency. Furthermore, image maps cannot be read by the blind and cannot be indexed by search engines.

  When an image map is appropriate, construct it following the previous recommendations on how to use graphics (small file size, useful ALT text, etc.). In addition, provide an alternative means of accessing the links in the

image map. A typical alternative is a list or menu of text links immediately below the map, or use text (or clear text) at the end of the page for indexing and for blind users.

Design tip: Don't use an image map that takes up the entire page.

Design tip: If a search engine can't follow links from your image map, it won't be able to index your site.

2.6 Multimedia

Examples of multimedia on the web include animations, audio, video, and three-dimensional models. The features are useful if done well. Keep in mind that most multimedia is slow to download and requires special viewers, so some users will not be able to enjoy these features.

2.6.1 Audio

Never use background audio on a web site.

- From RealMedia to MP3, audio is coming into its own as a web communications media. However make sure that the audio is used appropriately. Most users do not want background music unexpectedly blaring when they visit a site. Furthermore, unnecessary audio is a waste of bandwidth.

  Short audio clips and continuous streaming audio can be used to present audible content, if the user is aware of this feature and has the proper software. Much bandwidth is required, so distribute audio judiciously. Identify the type of content to be provided and the size of the clip, if known. Provide a written transcript for archived audio.

2.6.2 Video

Use short video clips and continuous streaming video to present visual content, if the user has the proper viewing software. A lot of bandwidth is required to support video on the web, so distribute video judiciously. Identify the content to be provided and the size of the clip, if known. Provide a written transcript or still images when possible.

RealMedia, QuickTime files, and streaming media broadcasts do take time to download, particularly on a slow modem connection. Both formats will play while downloading, which may improve the experience. Be aware that interruptions due to network congestion may occur while receiving data.

Design tip: When using video, warn the user of the limitations of slow modems, keep file size to a minimum, and provide links to a location to download the viewer if needed.

2.6.3 Three-Dimensional Models

Three-dimensional (3-D) content is a powerful tool for displaying data and objects that are not easily viewed in 2-D. VRML (Virtual Reality Modeling Language) or QTVR (QuickTime Virtual Reality) 3-D formats give the user the opportunity to interact with and view an object from all sides. However, the technology is still immature. Not all users have the viewing software or the adequate processing power necessary to render the models. These tools should be used selectively as a supplement to traditional content based on text and images.

Design tip: A simple animation that shows a solid object moving (for example, a rotating earth) may be more compactly represented in VRML (as object plus code to make it rotate) rather than as a movie.

2.6.4 Animations

Animations are short sequences of images that play either once or loop repeatedly, usually without sound. Typical formats include animated GIF (GIF89a), MPEG and QuickTime.

Flash is another form of animation. Flash sends streaming vector files that are much smaller in size than GIF files and can also provide MP3 quality audio capabilities. However, Flash does require a plug-in. Rather than require the user to download the plug-in to view the page, use JavaScript to test the user's browser for the correct plug-in. If the plug-in is not available, use an HTML page, with a static image. To reduce the plug-in issue, it is best to relegate large Flash animations or interfaces to one section of your site. Also, because users don't have the most recent versions of the plug-in, save your Flash with audio files to the earlier 3.0 version. Otherwise, users may not be able to use the sound files.

Animated GIFs can be inserted like a normal image into a web page. Only the first frame will be shown if the browser cannot perform the animation. GIF89a images are often used in banner advertisements. Animated GIFs are not recommended because the endless looping tends to become annoying to the user and because of the large file size required for multiple frames. However, if a brief animation is required, GIF89a is a suitable format.

MPEG, RealPlayer and QuickTime formats play either in a separate viewer or directly in the browser window. They are better than GIF89a for animations and for video clips that are more than a few frames long. Use them as needed for content. Always warn the user about the file size of the download. Whenever possible provide one or more still frames as preview images. This lets the user decide whether to download. At a minimum you should provide a text description.

QuickTime works better on Macintosh and MPEG works better on PC and UNIX machines, and RealPlayer has more cross-platform compatibility. Also be aware that this is not smooth video quality and movements may be jerky.

2.7 Cross-Browser Compatibility

Ensure that a site is usable on a variety of browsers.

- Users have browsers both modern and ancient, ranging from Macs, PCs, and UNIX boxes to VMS machines. Some users will choose not to load images; blind users, for example, use text-to-speech readers; power users have big monitors; wireless warriors use Palm Pilots and cellular modems; web TV watchers surf on their couches. A page that looks great on one screen, with a particular browser, using a specified font size, may be unattractive or unreadable on another screen. Therefore **you must** test pages using as many different viewing configurations as possible.

  Cross-platform testing can be performed manually, and can be automated in part by third party utilities. After a while it becomes clear what constitutes cross-platform HTML and explicit testing can be done less frequently.

Design Sites for All Users

NASA disseminates information to the broadest possible audience. Therefore to reach a broad audience, avoid requiring specific browsers. Test sites for browser compatibility instead of dictating a particular browser that requires the public to view the site with browser ABC.

Like all federal agencies, NASA must comply with the Americans with Disabilities Act. This compliance requires that web sites be accessible to the blind or vision impaired. Either every page must be accessible (readable by a text-to-speech converter, for example), or a parallel set of accessible pages must be provided.

Design Tip: If your site is for public viewing, then you should test the site on at

least than three platforms (Mac, PC, UNIX) using at least two browser brands per platform. A brand of browser such as Netscape or Internet Explorer will display the same site differently on different platforms. Also, be sure to modify the browser settings to turn off image loading for navigation testing.

2.7.1 Web Page Printability

Print capability should be considered when the site is designed. For example, dark color backgrounds do not print out clearly on some printers. Fixed width tables and graphics may be clipped off when finished.

For a printable page, keep page elements within a 500 to 600 pixel width. Another common practice is to offer a “printer friendly” version of pages, such as fact sheets, which the user may wish to print out. This HTML version does not have dark background colors or non-standard link colors and prints out neatly on an 8 1/2” x 11” sheet, without unwanted clipping by the printer. PDF versions are also “printer friendly.”

3.0 Site Accountability

NASA web documents are subject to the same policies governing hardcopy documents. This chapter describes:

- Why NASA sites must adhere to government policies
- What policies apply to NASA web sites
- Who is responsible for ensuring that a site conforms to these policies
- What content can and cannot be included in a public-viewable site
- How to respond to internal and external users' inquiries on policies

3.1 Accountability Defined

NASA web documents are subject to the same accountability and dissemination policies and guidelines governing hardcopy documents.

Site accountability refers to the person or persons responsible and liable for the content of a NASA web page or site. The "NASA Public Affairs Internet Policy" states:

> “NASA employees should follow the same general guidelines for posting information to the Internet that now guide their actions with other media, including television, radio, magazines and newspapers.”

Documents created or residing on government machines and intended for public viewing are considered official government records and publications. Therefore, these documents are subject to the same legal and regulatory guidelines and policies applied to hardcopy materials and publications. Therefore, anyone who publishes web pages on NASA servers is accountable for the content of those sites. This section strives to help the reader understand the significance involved in being the “accountable party” and to heighten the NASA web developer’s sense of duty with regard to publishing information on the web.

3.2 Appropriate Content for NASA Sites

NASA is candid in dealings with the public, and a similar stance is taken with the Internet. The "NASA Public Affairs Internet Policy" states,

> “From its origins in 1958 to the present day, NASA has pursued a policy of openness and accessibility. The Agency’s business is conducted and explained in full public view, and the results of NASA’s activities are considered a resource to be made as widely available as possible. With the advent of the Internet, a medium NASA helped create, a powerful new communications tool is available for explaining NASA’s programs to the American people, and for making many of the

results of those efforts widely available electronically. Information is to be freely shared across the Internet, but only information that is appropriate and cleared for public release."

Documents that are in development or are not to be viewed by the public should be placed on web servers that have restricted access. No document in a draft or development state should be viewable by users. Likewise, documents not cleared for public release should be separated from those posted on a public web site. In terms of content, all documents, whether public or private, shall be directly related to the official responsibilities of individuals and organizations fulfilling their assigned missions and charters consistent with their individual Center, NASA, and the federal government.

If NASA content will be included or referenced on a non-NASA site (either by directly including it in the web page or by including a link to another site), make sure that the content is pertinent to your organization, and that the resulting page content is appropriate and endorsed by a government agency.

3.3 Policies and Guidelines for Information Dissemination

The Office of Management and Budget (OMB) through its draft guidelines has clearly indicated that federal web sites fall under the existing policies governing federal information dissemination. (Refer to OMB Circular A-130) (http://www.whitehouse.gov/OMB/circulars/a130/a130.html)

The NASA Internet Usage Policy, Executive Notice 02-95 (ftp://ftp.hq.nasa.gov/pub/cio-office/Exc-Notice/en2-95.doc), establishes NASA's position for electronic publication and dissemination of NASA information for public access over the Internet is applicable to any information stored on the "nasa.gov" domain.

Government records are subject to a Freedom f Information Act (FOIA) (http://www.epic.org/open_gov/foia/us_foia_act.html) request. "Those seeking information are no longer required to show a need for information. Instead, the 'need to know' standard has been replaced by a 'right to know' doctrine. The government now has to justify the need for secrecy." In 1996 Congress amended this act "to provide for public access to information in an electronic format, and for other purposes" through The Electronic Freedom of Information Act Amendments of 1996 (EFOIA) (http://www.epic.org/open_gov/efoia.html).

Bear in mind that internal policies particular to the needs and requirements of individual Centers may also apply. Be certain that any information that is being made available complies with the guidelines and regulations of the organization and the NASA Center hosting the web site.

Design Tip: "Better safe than sorry" applies when posting documents to the web. If you are not sure if a document should be viewable to your intended audience, for any reason, then have appropriate personnel review the document before posting it to your site.

3.4 Identify the Site's Accountable Person or Organization

NASA web sites should contain the name of the person or organization responsible for the site's content and operation.

In Executive Notice 02-95, Internet Usage Policy (ftp://ftp.hq.nasa.gov/pub/cio-office/Exc-Notice/en2-95.doc), Section 1.3, Paragraph 1 states " …there must be clear accountability for the accuracy and appropriateness of the information displayed." Therefore all government sites should display the name of a point of contact (POC) who can explain what is contained within the site.

What this policy statement does not specify is how that accountability is to be displayed: Should the accountable party be a person or an organization? and what process should be in place so that users who pose questions to the site's accountable party will receive a quality response in a timely manner? Feedback from site users is important because it keeps the site's creators in touch with the needs of the audience and provides users with a way to get information to the site's accountable POC. The following sections provide suggestions for completing these tasks.

3.4.1 Identify Site Curators

The site curator is responsible for developing, deploying, and maintaining the web site. Place the name of the site curator at the bottom of the site's home page at the very minimum, and preferably at the bottom of all site pages.

If problems arise such as links that do not work, images that do not display, or the site cannot be found, the site curator will be contacted and is responsible for correcting the problem. The curator can be a civil servant(s) or a contractor(s).

It is best practice to adhere to a naming convention to identify the site curator, and maintain this convention throughout the site. Suggested naming conventions:

Site Curator: John Q. Public (john.q.public@nasa.gov)

Site Construction: John Q. Public (john.q.public@nasa.gov)

Where possible and/or applicable, consider using an HTML form to generate the e-mail for the user. A form has the advantage of automatically encoding critical information (such

as the URL of the page in question) and allowing the user to contact the site curator regardless of whether or not the curator has an e-mail account. For more information on using forms for feedback, see Section 3.4.3.

Design Tip: Include a "mailto: " link to the site's curator. You may want to create an e-mail alias (i.e., webteam@nasa.gov) instead of using a specific individual's e-mail account. E-mail aliases help hide the "e-mail identity" of the site's POC and allows for the adding or removing of multiple POCs by simply modifying the site's mailto: link.

Whether using an e-mail alias or a specific name, the full electronic mail address should also be displayed for those users who cannot or have not configured the electronic mail capability within their browser application. In addition, adding a default subject line to the mailto: will allow the site curator receiving the e-mail to easily determine which web page is referenced.

3.4.2 Identify the NASA Official Responsible for Site Content

Site officials must be civil servants and are responsible for the content of the site. Their role is to approve all information to be posted to the site and to review that information once it has been posted. **Place the name of the NASA Official at the bottom of the site's home page, at the very minimum, and optimally at the bottom of all site pages.**

Before posting content, use a staging server where pages can be approved before publishing the content. Once the site is viewable, conduct periodic reviews of the site's content to the site to ensure that information is relevant, accurate, and up-to-date.

As suggested in Section 3.4.1, follow a naming convention to identify the NASA Official, and maintain this convention throughout the site. Suggested naming conventions might be:

NASA Official Responsible for Content: John Q. Public (john.q.public@nasa.gov)

Page Content: John Q. Public (john.q.public@nasa.gov)

NASA Point Of Contact: John Q. Public (john.q.public@nasa.gov)

Where possible and/or applicable, consider using an HTML form to generate an e-mail response for the user. A form has the advantage of automatically encoding critical information (such as the URL of the page in question) and allowing the user to contact the site's NASA POC, regardless of whether or not this contact has an e-mail account. For more information on using forms for feedback, see Section 3.4.3 below.

Design Tip: It is best practice to include a "mailto: " link to the site's responsible point of contact. You can create an e-mail alias (i.e., nasa_poc@nasa.gov) instead of using an

individual's e-mail account. E-mail aliases help hide the "e-mail identity" of the site's POC and allows for the adding or removing of multiple POCs with modification to the site's mailto: link.

3.4.3 Timely Response to Site Inquiries

All site inquiries should be answered in a timely and accurate manner, and to the best of a person's or organization's ability.

According to the *Yale Style Guide* (http://info.med.yale.edu/caim/manual/contents.html)

> "Feedback also means being prepared to respond to your user's inquiries and comments…. Planning for this kind of ongoing relationship with the users of your site is vital to the long-term success of the enterprise."

Respond to each e-mail inquiry with a personal response that acknowledges receipt of each inquiry and provides appropriate feedback. To provide users with quick and personal responses, use a feedback form to:

- Help filter inquiries. Use radio buttons to ask the users if they expect a response. Some users just want to send the site curator comments.
- Provide an immediate response to all users through a "canned e-mail" that is automatically sent by the script processing the form's information.
- Provide users with the capability to select a subject of interest. The users' selections are sent electronically to personnel within an organization who respond quickly. This process distributes the responsibility of responding to inquiries to many people and organizations rather than placing the burden on one person.
- Redirect users to a Frequently Asked Questions site. Here they may find answers to commonly asked questions.

  Tip: The people sending e-mail to your site deserve a quick and polite response. They don't know you have been asked the same question ten times this week, nor that you handle a thousand other web activities. No matter what your time constraints, always respond quickly and politely. Your response is a reflection on NASA and at all times must be clear, concise and civil.

3.5 Hit Reporting

Some sites require monthly NASA reports that may include these types of statistics:

- Number of total hits

- Quantity of data delivered
- Unique network IP addresses served

These statistics provide information that shows user traffic patterns including:

- The most popular pages on a site
- Length of time a user spends on an individual page
- Number of return visits to a particular site section
- The path a user took through the information

To obtain user information, find a web-hosting service provider that offers parsed log files that can be captured and analyzed on a regular basis. These tools are commercially available and offer a variety of useful reports on user preferences.

Another approach is using web-hosting log reports that at minimum provide information on the basic “big three” web statistics: number of total hits, quantity of data delivered; and unique network addresses served.

Detailed user information helps webmasters revise a site to reflect user preferences. Webmasters can use web statistics to restructure site information to facilitate traffic patterns.

4.0 Site Security

NASA depends upon the WWW to disseminate information both internally to its employees and externally to the public. To maintain the integrity of NASA's information and systems, Agency information must be adequately protected yet readily available to all intended audiences.

What are the elements that will make a web server secure? There are 5 critical security elements. Their order of importance will vary for each site, but the following criteria in this order are common to most sites:

1. Availability or "Can people get to my server?"
2. Privacy or "Is anyone eavesdropping?"
3. Authentication or "Who are you?"
4. Authorization or "Who is allowed to do what?"
5. Non-repudiation or "I know what you did, and you can't deny it."

4.1 Availability

Availability simply means "can people get to my web site?" When users come to your site, they expect that link to work right—every time! Otherwise you've lost the user.

Tip: Start with hardware. Use the best CPU, as much RAM as you can, the most disk space, and the fastest network connection you can get.

Things to remember:

1. Hardware is cheap, compared to the labor for a workaround.
2. Hardware will be obsolete next year anyway.
3. Load will only increase, never decrease.

4.1.1 Monitor Resource Usage

Monitor your resource usage. Is the CPU idle or overworked? Is the network swamped? Are your disks filling with log files? Keep statistics on the usage of each resource over time so you notice what you're running out of, and at what rate. For starters, on UNIX try the **vmstat**, **netstat**, or **sar** commands. On Windows NT, right-click on the task bar, choose "Task Manager," and look at the "Performance" and "Processes" tabs.

Tip: Use load balancing. If you've got more traffic than one machine can handle, use a pool of smaller, cheaper machines. Mirror the data so it's the same on all of them, or use a caching shared file system like AFS to replicate the data across machines.

4.2 Privacy

Privacy is gauged by how easy it is for someone to eavesdrop on what the user and server send back and forth. You wouldn't want to send a credit card number to a server if that information could be read as it crossed the network. Likewise you don't want to download a sensitive or restricted document that might be viewable to the public.

4.2.1 Use Secure Sockets Layer (SSL) Technology to Protect Information

One way to increase privacy in a web transaction is to use Secure Sockets Layer (SSL) technology. There are many commercial sites that use SSL. This technology is being used whenever URLs appear with "https" (HyperText Transfer Protocol, Secure) instead of the "http," which is used for unsecure sites. The https uses SSL to encrypt the traffic between a web browser and a web server, even if the information can be seen by the user, it does not appear as legible text.

SSL protects the information the user and server exchange. But what if the connection setup is intercepted? For example, you try to connect to microsoft.com, but your connection is hijacked by Dr. Evil, creating a private connection to Dr. Evil. Now all your information is sent to this site.

To avoid information from being hijacked, use server certificates. Server certificates solve the issue of knowing who you're dealing with; for details, see Section 4.3, Authentication.

Tip: Virtual Private Networks (VPNs) are another way to increase privacy., VPNs encrypt all traffic, not just web connections. Because VPNs encrypt traffic between networks instead of individual machines, they are more applicable for intranets (connections within a company or Center).

4.3 Authentication

Authentication means identifying both the user and server to ensure proper usage. A user wants to be sure a server is authentic before sending personal information like a credit card number. How do you really know that you are really connected to amazon.com and not some imposter?

4.3.1 Server Certificates

Server certificates (digital letters of introduction) are installed on web servers. A Certificate Authority (CA) issues a certificate and signs it. When you go to a web site, the server's certificate is presented to your browser that authenticates the CA.

There are several dozen CAs that most web browsers can authenticate. For instance, Netscape browsers trust the Netscape.com CA, and MS Internet Explorer browsers trust the Microsoft.com CA. VeriSign and Entrust are trusted companies as well. These certificates are used in SSL, described in Section 4.2, Privacy.

Authenticating a user often requires a name and a password, though for some information it's enough to check that a web request is coming from a machine in a specific domain like the *.nasa.gov domain.

Although not all-inclusive, here are three examples of web pages with different authentication needs:

1. *A press release publishing science data results.* The authentication decision here is moot, since the information is written for public release.

2. *An internal mission status report.* This information can be distributed throughout NASA, but not to the public. Checking that the request came from *.nasa.gov domain may be adequate.

3. *A draft of a procurement RFP*. This information is highly sensitive and only procurement team members should see it. You need some way to prove the requester is part of this team.

Even in the first example, the decision not to check is still a decision. This is how most web servers act by default, yet many webmasters are still surprised when the information published can be read by 15-year-old Canadians.

4.3.2 Domain Name System Registration

The second example, checking the requester's IP address or domain name, is commonly used even though it is a weak solution. This decision is better than nothing, but basically keeps out the honest people and not the serious hacker.

Tip: Decide what set of addresses to give access. This can be by domain name, like the following:

```
        *.jpl.nasa.gov
or      *.nasa.gov
```

It can also be by the actual IP number or subnet (range of IP numbers) like the following:

```
        137.78.81.46
or      137.79.*
```

If you use the name, each requester's machine must be properly registered in the Domain Name System (DNS) Registration.

Tip: You can check whether an IP address is registered by giving the following command from any UNIX or Windows command prompt:

nslookup 198.116.142.34

This command will either tell you the machine's hostname or will report that the machine is not in the DNS; in the latter case the requester will have to fix the DNS.

Tip: Domain authentication is risky at best, so use it only if you really don't care whether the information is published publicly. Or, combine it with the stronger authentication, as outlined below.

4.3.3 How to Authenticate Users

In the third case, you really need to know who someone is. Greek soldiers guarding their camps and challenging strangers for "the password" had the same problem. All approaches and technologies boil down to the same thing: prove that you know something no one else could know, or that you possess some token (key, artifact, religious icon, etc.) no one else could have.

There are several ways to challenge and authenticate users attempting to access your site.

- *Username and Password*. Most web servers allow you to create usernames and assign them passwords. Servers can also be configured to authenticate using other systems that manage passwords, like kerberos realms and NT domains. Train users to select good passwords, and configure your system to reject bad ones. A poor password can be easily guessed. Reject words in the dictionary, common names, etc. Make passwords expire every 6 months or so to force users to change them, and keep a history of previous passwords so the user doesn't repeat them.

- *Digital Signatures*. Use LDAP (Lightweight Directory Access Protocol) authentication, or the NASA-wide PKI (Public Key Infrastructure), which is being developed by Ames Research Center. Contact the Ames team to find out how to use LDAP and PKI.

- *Tokens*. These are physical objects like a Secure ID, smartcard, floppy disk, fingerprint, or retinal scan that are used for authentication.

It is prudent not to rely on just one type of security measure. Passwords are often transmitted “in the clear” from browser to server, thus leaving passwords open to interception by hackers. It is astounding how many places software writers stash passwords for clear-text transmission. Therefore, in conjunction with passwords, you also may want to use SSL to establish an encrypted connection.

4.4 Authorization

Authorization is also called “Access Control” and determines the privileges a user has on a system. It is important to identify users to determine their level of access.

There are various ways to configure certain types of authorization on a web server. For example, web server configuration files like httpd.conf (Apache) can be set up by the administrator to allow or deny access to an entire web site. These files can also restrict the placement of a CGI (Common Gateway Interface) script to reduce the risk of a dangerous program from being run on your server; in addition, the robots.txt file resides in the web server’s document root and is used by automated search engines to determine whether to index pages on the web server.

Directory configuration files like “.htaccess” (Apache, NCSA) and “nsconfig” (Netscape server) allow you to specify what pages, directories, etc. can be viewed by a particular IP address, domain, or name/password. These files are usually more specific than a site-wide file like httpd.conf.

4.4.1 Securing Servers

Because web servers are the most visible targets for intruder break-ins, they need to be as secure as possible. Therefore you must minimize the means of unauthorized access. Here are some tips for securing servers:

- When a machine runs a web server, all other services should be disabled, including e-mail, finger, ping, telnet, and ftp (except maybe anonymous ftp, carefully configured).

- Web administrators should schedule periodic security scans by a trusted third party in order to identify security weaknesses as well as look for suspicious behavior like an attempt to access the old "ph" CGI program.

- Minimize your overall system risk by minimizing what the web server can do. Never run the web server as "root" or "administrator." Always create a new user account to run the server process, with no other privileges on the machine.

- Use a shared file system like AFS or NFS and give the web server "Read Only" access, or mount the data disk "Read Only," or if the content rarely changes, burn it into a CD and serve the site from there. When the web server is compromised, an intruder still won't have the ability to modify data.

Because employees come and go, authorization configurations must be changed as personnel changes. A new developer or manager must have access to important information that keeps the system running smoothly. Utilizing role-based authorization may be a good practice if there is frequent turnover of those requiring access to your site. Role-based authorization would list a group like "shuttle_pilots" on an ACL (Access Control List), rather than listing specific people. The group contains the individual users who have that job or work on a specific team. New personnel are added to the group and automatically have access to whatever the group has been granted.

4.5 Non-Repudiation

According to Webster's Dictionary, one definition of "repudiate" means to reject as untrue or unjust. Basically, it is very important, in terms of site security, to keep good logs of your web server. Proper logs can provide sufficient proof of a user's activities by logging the actions of that user. For the most part, this is usually unnecessary unless there has been some questionable activity on the web server. Proper logging provides a means of non-repudiation for security officials in the event that someone has acted maliciously against your web server.

Because log files can be tampered with or erased, a best practice is to keep all logs on another machine, or log files could be written to write-once media like a "write-only" compact disk.

It is important to note, however, that there are certain guidelines on gathering and retaining information. Please be aware of the Children's Online Privacy Protection Act of 1998 (COPPA) (http://www.ftc.gov/opa/1999/9910/childfinal.htm) and the NASA Chief Information Office (CIO) policy on retention of web logs found under the NASA Records Retention Schedule (http://www.sti.nasa.gov/nasarrs). The CIO web site points to item 13A for guidance on the retention of web logs. Item 13A is found on page 2-16 of the NASA Records Retention Schedule PDF file.

Non-repudiation relies on authentication to know who someone is. Obviously if authentication is weak then so is your ability to know who's doing what. You always know what IP address was in each request because that's the destination where the information is supposed to be returned.

4.5.1 Non-Repudiation Tips

- When possible and practical, log everything the web server does. Every server can keep a log of every access attempt sent to it, as well as every error it may have had. Keep all of the logs for a minimum of a year or as close to forever as you can.

- After time, log files can take up valuable server space. Develop a program or procedure to compress the log files periodically. The logs compress at about 10:1, meaning that a 10 MB log file only takes 1 MB if compressed. Save the older log entries, compressed or not, somewhere else like to tape, JAZ or ZIP files or on another disk drive.

4.6 Current NASA Policy on Protecting Web Servers

Procedures and guidelines for determining the appropriate measures necessary for protecting information and for determining protective controls required for computer systems that host web pages can be found in the NASA Policy Guideline 2810, "Procedures and Guideline for the Security of Information Technology."

You should review NASA-STD 2813, "Firewall Strategy, Architecture, Standards and Products."

4.7 Who Makes the Determination of What Information Should Be Secured?

The "data owners" are responsible for determining the dissemination requirements for their information. These dissemination requirements drive the security controls required for the information and the web sites.

The author of the document (data owner) must make the first determination. When the data owner has determined the dissemination requirements for the information contained on a web site, the data owner must then place the document on a web server that has security controls in place that can ensure the appropriate dissemination requirements.

Tip: Do not assume that all documents on a site have the same security requirements. Site document composition varies and sometimes portions of a site must be maintained more securely than others. Evaluate each document independently.

4.8 What If a Site Has a Security Breach?

Immediately report web security breaches to your Center Information Technology (IT) Security Office. Do not take immediate corrective action yourself because evidence of the breach may be destroyed.

4.9 Who to Contact

Ames Research Center (ARC) is the Principal Center for Information Technology Security. ARC maintains a list of the Center Information Technology Security Offices and outlines the steps for reporting suspected security breaches and incidents.

NASIRC (NASA Automated System Incident Response Capability) maintains a checklist and a FAQs site to help secure a system and answer operating system specific questions.

Tip: Use products that automate monitoring of web servers and the underlying operating system and provide automated notification should a security breach occur.

4.10 Case Study: JPL

JPL's main internal web server runs the Netscape Enterprise server. This server reads data out of AFS, a shared file system that is integrated with a kerberos4 authentication realm. Pages get written by authors to AFS file servers, and then are read by the web server from AFS. Although authors can change their content at any time, they do not have access to the web servers. The web server has permission to read only some directories on the file servers, but not (usually) to write or change anything.

Directories are protected with "nsconfig" files that restrict access to JPL IP subnets by default, though users can change the "nsconfig" to allow any kind of access from world-readable to a single IP address. When restricting content to a specific user name, the AFS name/password is used so authors don't have to create accounts themselves.

A robots.txt file at the document root forbids search engines access on the server, except those engines that JPL runs and are listed by name in robots.txt.

References

· NASA Policy Guideline 2810, "Procedures and Guideline for the Security of Information Technology"
http://nodis.gsfc.nasa.gov/Library/Directives/NASA-WIDE/Procedures/Legal_Policies/N_PG_2810_1/contents.html

· NASA-STD 2813, "Firewall Strategy, Architecture, Standards and Products"
http://www.hq.nasa.gov/office/cio/standards/2813.pdf

· NASA Principal Center for Information Technology Security
http://computer-security.nasa.gov/

- NASIRC (NASA Automated System Incident Response Capability)
  http://nasirc.nasa.gov/

- World Wide Web Consortium Security Resources
  http://www.w3.org/Security/

- General Services Administration Compilation of Some Key Policies from various government agencies - Top Ones for New Webmasters
  http://www.itpolicy.gsa.gov/eagency/overviewlevel3/webmasterpolicies.htm

- Federal Computer Incident Response Capability General Guidance for Security a Public Web Site
  http://www.fedcirc.gov/guideindex.html

- Standard for Robot Exclusion
  http://info.webcrawler.com/mak/projects/robots/norobots.html

- Certificate Authorities
  http://www.verisign.com/

- PKI (Public Key Infrastructure)
  http://www.entrust.com/

- AFS shared file system
  http://www.transarc.com/Support/afs/

- IBM WebSphere load balancing
  http://www.ibm.com/software/webservers/

5.0 Accessibility

> "The principles of universality of access irrespective of hardware or software platform, network infrastructure, language, culture, geographical location, or physical or mental impairment are core values in web design…."
>
> Tim Berners-Lee, *Web Architecture From 50,000 Feet*
> http://www.w3.org/DesignIssues/Architecture.html

The World Wide Web empowers people by giving them interoperable, networked access to a vast array of information. NASA endorses the *Web Content Accessibility Guidelines* (http://www.w3.org/TR/1999/WAI-WEBCONTENT-19990505) of the World Wide Web Consortium whose purpose is to extend the accessibility of the web to people with disabilities. The guidelines extend web accessibility across a variety of user agents (mobile phones, handheld computers, automobiles, and computers). These guidelines also enable access under a variety of conditions, such as in abnormal lighting conditions or situations where manual operation is not possible. Adoption of the guidelines facilitates finding information on the web.

Some of these guidelines are currently under review for adoption as standards for Section 508 (http://www.section508.gov/) of the Rehabilitation Act of 1973 as amended in 1998. The standards, as proposed by the United States Architectural and Transportation Barriers Compliance Board, are scheduled for inclusion of the Federal Acquisition Registry in August 2000.

These standards, along with references to the appropriate sections of the "Web Content Accessibility Guidelines," are excerpted below:

1) A text equivalent for every non-text element shall be provided via "alt" (alternative text attribute), "longdesc" (long description tag) or in an element content.
   - W3C References
   - Checkpoint 1.1
   - Technique 4.7.1

2) Web pages shall be designed so that all information required for navigation or meaning is not dependent on the ability to identify specific colors.
   - W3C References
   - Checkpoint 2.1

- Technique 5.5

3) Changes in the natural language (e.g., English to French) of a document's text and any text equivalents shall be clearly identified.
   - W3C References
   - Checkpoint 4.1
   - Technique 4.2

4) Documents shall be organized so they are readable without requiring an associated style sheet.
   - W3C References
   - Checkpoint 6.1
   - Technique 5.1

5) Web pages shall update equivalents for dynamic content whenever the dynamic content changes.
   - W3C References
   - Checkpoint 6.2
   - Technique 4.10.4

6) Redundant text links shall be provided for each active region of a server-side image map.
   - W3C References
   - Checkpoint 1.2
   - Technique 4.7.6

7) Client--side image maps shall be used whenever possible in place of server-side image maps.
   - W3C References
   - Checkpoint 9.1
   - Technique 4.7.5

8) Data tables shall provide identification of row and column headers.

- W3C References
- Checkpoint 5.1
- Technique 4.5.1

9) Markup shall be used to associate data cells and header cells for data tables that have two or more logical levels of row or column headers.

- W3C References
- Checkpoint 5.2
- Technique 4.5.1

10) Frames shall be titled with text that facilitates frame identification and navigation.

- W3C References
- Checkpoint 12.1
- Technique 4.10.1

11) Pages shall be usable when scripts, Applets, or other programmatic objects are turned off or are not supported, or shall provide equivalent information on an alternative accessible page.

- W3C References
- Checkpoint 6.3
- Technique 4.12.1

12) Equivalent alternatives for any multimedia presentation shall be synchronized with the presentation.

- W3C References
- Checkpoint 1.4
- Technique 4.8.3 & 4.9

6.0 Web Marketing

Building a web site without harnessing the power of search engines is like posting a billboard in your basement. Only your best friends will know it's there! Due to the sheer number of web pages competing for attention on the Internet, you need to make sure that your web site is easy to find. Users may select a search engine based on the proximity of a search button, the ease of use, the quality of the results they receive, or its popularity. Since most people depend on search engines to find information, knowing how to attract search engines to your web site is a skill you should acquire.

6.1 Overview of Search Engines

The ultimate goal of every web site is to be included in the first 10 search results. But how do you accomplish this?

Generally speaking search engines read and memorize the words on a web page. If the word that the "searcher" enters into the query is not on the page, no match results and the page will not be listed. However each search engine uses different techniques for its searches. Search engines may look for titles, content, keywords and descriptions (as included in HTML "META tags") or some combination of the above.

Once the content has been developed, the search engines use algorithms to rank the pages by generating a Confidence or Relevance Factor that is displayed as a percentage or points with the search results.

The Relevance Factor is determined by how many times a word is used in a document and how it is used. Ranking also may be determined by a site reviewer or by the number of web sites that link to your web site, indicating its credibility or popularity.

6.2 How Search Engines Acquire Content

To maintain a competitive edge, commercial search engines companies closely guard how they acquire and rank their content. No two search engines select their content in the same way, and they change their algorithms often. (An algorithm is a set of rules or commands.) Since these algorithms are not published, developers have difficulty creating web pages that will attract a particular search engine. However, most search engines have a few common traits, such as how spiders (see Section 6.2.1) gather content. These issues will be addressed in this section.

For additional reading on this topic, see:

- "Search Engine Secrets"

http://www.pegasoweb.com/freerep.htm

- "Sending Search-Engine Traffic to Your Site"
  http://hotwired.lycos.com/webmonkey/99/31/index1a.html?tw=frontdoor

6.2.1 Spiders

Most search engines acquire content by using spider programs that "crawl" your web site and index its contents. The search engines keep a "seed list" of URLs that are used to start the "spider" program on its way to your web site. While the guiding algorithms for each spider are configured differently, all spiders capture specific information from the pages. Some spiders capture *the title* and *first 1,000 characters* of content. Others capture the *title plus the description* and keyword *META tags.* Some capture a combination of all of these.

6.2.2 Submitting Content for Marketing Your Site

6.2.2.1 Indexing your site

Large or small, an organization must find the means to be listed by the major search engines and directories—one of the quickest and most cost-effective means for drawing traffic to a web site. You can get your web site listed by submitting information about the web site to a search engine (or search engine listing service) requesting that your site be indexed. Although many listing services are free, some charge a fee. Once you've submitted your request, it can take several weeks or even months for your site to be indexed. If you can't wait, some search engines offer "express" indexing for an additional fee. Express indexing moves your web site to the head of the indexing queue so your site will be indexed sooner.

6.2.2.2 Directories

You may also submit information to a directory site that categorizes the World Wide Web based solely on input that has been submitted. The site's representative submits information such as a site title, description, keywords, and URL. This information is then categorized by topic. In this case, a search engine never accesses the actual web site's content. When someone searches with a keyword, it is referenced against the database of sites for the closest match.

For more tips on submitting content to commercial search engines, see Section 6.5.

6.2.3 Editorial Selection

The editorial selection process uses human editors who assess site content for searchable words and concepts. Yahoo uses editorial selection. Although it is a more personalized approach, the process does have problems. For example, editors often use inconsistent selection methodology. Also, editorial selection is not easily scalable and is a labor-intensive process that often causes lengthy backlogs.

6.2.4 Link Analysis

Link analysis is one of the newer ranking methods that analyzes the number and quality of links to and from your site. If many sites are linked to yours through a specific keyword, your site is considered to be a relevant site. In turn any sites that you link to also become more relevant. To find out more about link analysis, check out the following site:
Google (http://www.google.com) is a search engine that utilizes link analysis.

6.2.5 Popularity

A site's popularity is based on user traffic. Search engines monitor the search results people select most often for a given query. The links that appear to be the most popular are then listed in a sidebar called something like the "Top 10 Most Popular Sites." Similar to the books listed on the *NY Times* Best Seller list, the web "top-10" listings create a name recognition and web presence that bring in more users, while other lesser known sites are overlooked.

6.2.6 Paid Placement

If your site appears in the first 10 search results, you have prominent web presence. To gain this competitive edge, some search engines invite site owners to purchase or bid on key words that pertain to their web site. For example, if the site features sports-related information, the keyword "sports" might be purchased, ensuring that the site's URL would be included in the first set of search results for "sports." The purchase agreement may also guarantee that a clickable ad for the sports site will appear on the page with the results.

Some web sites identify paid-placement search results by including a notation such as "Cost to advertiser $0.41." However, this notation is not currently required, so most sites do not include it.

6.3 Developing Searchable Web Pages

6.3.1 Start with the Title

Titles help users quickly grasp the site's purpose. Good titles use descriptive, relevant keywords in the title for search purposes, making it easy for users to find your site. If the title doesn't reflect the contents of the page, a user may not bother to revisit your site.

Most search engines search by the title displayed at the top of the browser window; the same title that will appear in the user's bookmark. In the HTML code, the title is listed between the title tags, as shown here:

<TITLE>Your Title Goes Here</TITLE>

If you use frames—which NASA best practices discourages for design and disability reasons—make sure that all of the subordinate pages within the frames format also have titles. When search engines search individual pages, subordinate pages may be logged in the results, in addition to your main page.

6.3.2 Choosing Keywords

Choosing keywords for your META tags is an art—similar to cataloging information. To use keywords to your advantage, use specific and common terms for your site. If keywords are too unusual or are too difficult to spell, searcher may not include them in the search query. Some tips:

- Consider the terms the searcher would use.
- Make a thorough list of the keywords that best describes the information contained on the site.
- Give greater credibility to your keywords by frequently using keywords.

As keywords are used more frequently within the site, the Relevancy Factor grows higher. Consequently, the page is given priority in the search engine listings. However do this with restraint; do not overuse keywords; overuse can affect the document's readability. Some search engines will actually reject the keywords if they are repeated too often or spammed in the document. Some search engines only look for keywords in the body of the document.

6.3.2.1 Tips for Building Keywords List

- **Use Plurals**
  Plurals are easier to search than singular words because search engines capture both the singular and plural forms.

- **Vary Word Forms**
  Use the keywords in various forms, with hyphens and without, one word or two words, etc.

- **Consider the Common Misspellings People Make**
  If you are aware of how users commonly misspell words, you can use these variant spellings in your keywords. Certain terms like NASA are often misspelled. For example, searchers often type Nasa or Nassau.

- **Use Keyword Combinations**
  Some sites only look at the first 250 characters of a keyword list. To conserve your character count, create keyword combinations or "key terms." For example, if your site is about DS1 and Remote Agent, you could use "DS1 Remote Agent "as one key term, thus saving space for additional key words or terms.

- **Use Phrases**
  A recent study showed that users query phrases over 65% of the time. Some search engines even prefer phrases such as Ask Jeeves: http://www.askjeeves.com/

6.3.3 Use META Tags

In addition to title and keyword searches, many search engines target META tags placed within the HTML code. META tags are not visible to the user. Therefore, you can code several types of META tags within the HTML to make the web site more searchable; you can use description tags, keyword tags, and many more options.

Search engines index the entire page, regardless of whether descriptions or keywords are included in the <META> tags. The words in the <META> tags are indexed along with the rest of the document.

6.3.3.1 Make META Tags Descriptive

Some search engines use META tags as the source for the page description that accompanies the URL in the search results. Construct the description carefully. It should get the searcher's attention as well as describe the page. Use several of the most important keywords in the description and keep the description short. *Some search engines will not accept descriptions that exceed 25 words.*

To create a description for a web page, use the<META> tag within the <HEAD> element. The basic syntax is:

<META name="description" content="Write your description here">

Do not use HTML tags within the description or content part of the META tag.

The tag would be used in the HTML at the top of the web page in this manner:

<HEAD>
<TITLE>The Kids In The Hall Home Page</TITLE>
<META name="description" content="Home Page for The Kids In The Hall, Canada's funniest comedy group.">
</HEAD>

The information from the tags would be displayed in the search engine results as shown here:
The Kids In The Hall Home Page
Home Page for The Kids In The Hall, Canada's funniest comedy group.

6.3.3.2 Keyword META Tags

The search engine uses information in the META tags when it indexes the page. Using the <META> tag to add keywords to a page provides additional information about a page without interfering with the readability of the text.

Specifying keywords:

A second <META> tag can be used within the <HEAD> tag to specify keyword phrases that further describe the web page. The basic syntax is:

<META name="keywords" content="Write keywords here, in a comma separated list">

If paired with a description tag, the HTML at the top of the page might look like:

<HEAD>
<TITLE>Landing Venture Capital</TITLE>
<META name="description" content="Tips on obtaining venture capital for your hi-tech start-up.">
<META name="keywords" content="asset based lending, capital, enterprise funding, equity funding, expansion capital, funds for growth, growth financing, underwriting, venture investment">
</HEAD>

6.3.3.3 Use Robots META Tags

If you want to prevent a search engine from indexing a single page or pages of your site, place the Robots META tag in the HEAD tag of a page to deflect the robot or spider. The syntax for this tag is:

<HEAD>

<META NAME="ROBOTS" CONTENT="NOINDEX, NOFOLLOW">

</HEAD>

Not all search engines support the Robots META tag. Until the tag is more widely supported, you may have to restrict indexing of your site by another means.

6.3.3.4 Optional META Tags

Additional META tags that provide more information about a web site can be added as necessary. An industry consortium called the Dublin Core Metadata Initiative has developed a 15-element metadata set to facilitate the discovery of electronic resources by digital libraries, researchers, other agencies or specialized communities. Several of the Dublin Core elements are worth considering:

- **Creator (or Author):** An entity primarily responsible for creating the content of the resource
- **Publisher:** An entity responsible for making the resource available
- **Date:** A date associated with an event in the life cycle of the resource
- **Rights:** Information about rights held in and over the resource

More information and example on the Dublin Core Element Set is available at http://purl.oclc.org/dc/.

6.3.4 Development Techniques

6.3.4.1 Tricks to Use

6.3.4.1.1 Getting listed first—even if someone is using the same keywords

Some search engines will alphabetize their lists. If another competing site uses the same keywords, one way to get to the top of the list is to start the title with an "A". For example, the title "Wonderful Books" can be changed to "Absolutely Wonderful Books," which alphabetically put the title at the beginning of a search. Avoid using a single "A" as in "A Wonderful Book" since some search engines do not consider single characters when searching. Numbers are also good. Sometimes numbers are indexed before alphabetical listings.

6.3.4.1.2 Using the most-searched words

Be aware of what words are searched most often and try to include some of them as keywords. For example, by using the title “Dog Training for Men and Women” instead of just “Dog Training” the title would include two of the most frequently searched for words on the Internet (Men and Women).

6.3.4.2 Avoid Repeating Terms

One trick of the trade is to use the same term repeatedly. However this trick can undermine your search goals. Cramming a page with the same search term. Some search engines adopt a zero-tolerance policy for filling a web page with keywords that are repeated page after page; these engines reject the term if repeated too often.

6.3.4.3. Use “Invisible” Keywords

Using invisible keyword is a technique that when used properly increases the relevancy percentage used by search engines to put a page at the top of their list. To insert invisible keywords:

1. Go to the bottom of the web page and enter carriage returns until the last line of text is no longer visible on the page.
2. Using the smallest type size available on the browser, type the list of keywords
3. Change the font color of the keywords to match the background. This renders the words invisible to the reader, but not to the search engines. (Be aware that some search engines may consider this an attempt to stuff a page with too many keywords.)

6.3.4.4 Hide Keywords in Images

You can hide keywords in images to create better search results. When inserting an image into a document, you have an option to add text that will show if the image isn’t displayed. In some HTML authoring programs this is called an *alternate* or ALT tag description. The information in ALT tags can be read by the screen reader software used by the visually impaired, and can provide useful information about the images. (Again some search engines will consider this method stuffing a page with keywords.)

The actual HTML code is as follows:

<IMG SRC=“yourgraphic.gif” HEIGHT=30 WIDTH=60 ALT=“describe your graphic and then list your keywords”>

6.3.4.5 Hide Keywords in Comments

HTML programmers place comments in the source code to provide additional information or instructions about the code. A comment is placed in between angle brackets (e.g.: < >). For example:

<!-- rich money prosperity gold wealth -->.

Once again, be warned that the search engine may consider this an attempt to stuff a page with too many keywords.

6.4 Testing for Searchability

Testing a site for searchability involves looking at the results of various searches and analyzing the results as they relate to your web site. However, the process takes a long time to download each page, throw out invalid results, count the instances, and note the location of your URL. Luckily, there is a site that will make short work of this process.

Direct Hit at http://www.directhit.com is a popularity analysis search engine.

This site tracks how often a site is chosen from a given set of search results. Commercial products that will do keyword density analysis testing are also available.

6.4.1 Testing META Tag Effectiveness

To test for effectiveness, perform a search for home pages with similar content using several different search engines. See which pages show up at the top of the indices. Go to those pages and note how each has utilized the keywords used in the search. Also check the document source to see if META tags have been used for the description and keywords. Doing this on a few pages should provide some good keywords to add.

6.4.2 Testing Keywords

Similar to the technique above, keywords can also be tested. A developer should try individual keywords as well as word combinations on several search engines to see what types of home pages show up in the results. This test will give you an idea whether the keywords chosen are appropriate for the topic of your home page. If certain combinations of keywords seem to target the appropriate type of pages better than others, move these keywords to the beginning of your keyword listing to improve the your page's ranking by the search engine.

6.4.3 Testing Position Ranking

You can monitor how top search engines rank your web site by visiting sites that track these keywords. Here are a few of the sites:

- Position Agent
  http://www.positionagent.com

- Web Position
  http://www.webposition.com/

6.5 Tips for Submitting Web Pages to Search Engines

- Submit several key "entry point" pages of your site to search engines.

- Submit the Table of Contents or Sitemap.

- Prepare a brief site description and include it with the submission.

- Update your pages and resubmit for each iteration. Each time the "Last Updated" date changes, your site is eligible for new indexing.

- Consider submitting to a listing service that automatically submits your home page to several other search engines or directories for free.

- Submit your site to your Center or the NASA search engine.

- Establish reciprocal links with sites that are indexed by commercial search engines, such as your Center's home page, to be included in their indexing run.

- Become an information hub or a recognized source for good links, and the search engines will come to you.

When submitting a site, be patient. Due to volume and backlog some search engines can take 6 weeks or more to get a site listed. Others will show up in a couple of days. Do not be discouraged if a page does not show up immediately.

6.5.1 Where to Submit a Web Page

The first place to submit a web page is to the most popular search engines. Look for an instruction page that tells you how to submit your site. Keep an eye out for new search engines and add them to your submission list. The following is a list of places to submit web pages:

Alta Vista Search Engines
http://www.altavista.com/cgi-bin/query?pg=addurl

Webcrawler Search Engines
http://webcrawler.com/Help/GetListed/AddURLS.html

Lycos Search Engines
http://www.lycos.com/addasite.html

Infoseek Search Engines
http://www.infoseek.com/AddUrl?pg=DCaddurl.html

Excite Search Engines
http://www.excite.com/Info/add_url.html?a-m-m

HotBot Search Engines
http://www.hotbot.com/addurl.asp

Magellan Search Engines
http://magellan.excite.com/info/add_url

Yahoo Search Engines
http://www.yahoo.com/info/suggest/

6.5.2 Submitting Custom Pages

If you have the disk space to do it, create a separate web page that satisfies the requirements of each of the major search engines. The idea here is to take your "main" or "index" page and create several copies of it, e.g., maina.htm, mainb.htm, mainc.htm etc. With each copy, change the invisible components to satisfy particular search engines.

6.5.3 Submitting to Hundreds of Search Engines and Directories

Announcement and listing services allow you to submit a site simultaneously to many search engines and services that list web sites. Many of these services are free. The following is a list of some of these announcement or listing services:

Worldwide Broadcaster – A relatively new announcement service that allows submission to over 200 sites and search engines with only one submittal to Broadcaster!
http://www.broadcaster.co.uk/

Postmaster - Posts your site to 24 different places and search engines for free.

http://www.netcreations.com/postmaster/

Submit It! - Submits a site to 15 places and search engines for free.
http://free.submit-it.com/

Quick Launch - Submits a site to 15 directories and search engines for free.
http://www.qwiklaunch.com/

6.6 Secondary Marketing Tips

Building a web site without doing secondary marketing is like posting a billboard in your basement—few will visit your basement. Here are a few suggestions for using different types of media to augment the search engine's ability to bring visitors to your site.

- **Include your URL in print media**
  List it in press releases, newsletters; datasheets, posters, tradeshow exhibits, publications, business cards, T-shirts, and mugs, etc.

- **Publicize the URL on television**
  If the news media is doing interviews or covering an event at your Center, give them your URL as part of the news coverage.

- **Post notices about your site's content to newsgroups**
  Whenever you add useful non-promotional content to your web site, such as a "Frequently Asked Questions" file, post notices to relevant newsgroups, forums and mailing lists. Not only does this increase traffic to your web site, but it also prompts people to add links from their site to yours.

- **Start an e-mail newsletter.**
  Some say that showing up in your viewers' mailboxes every week or month with schedules or industry news is even more valuable than the web site itself. E-mail may be the most effective "push" technology to date. Always remember to invite your readers to check out your web site for more information.

- **Provide a unique service or capability**
  Add something unique to your site that no one else has such as a simulation, a space calculator or an image archive and then send out a news release about it.

- **Keep the web site support staff informed**
  Make sure that the web site's support staff is notified of your web site marketing activities. Your marketing efforts could cause an unexpected jump in traffic that could overwhelm your server.

- **Stay current with the latest marketing news and techniques**
  Read articles or visit discussion areas such as:

  - Search Engine Forum
    http://searchengineforums.com/bin/Ultimate.cgi

  - The Internet Marketing Center
    http://www.marketingtips.com/

  - Net B2B
    http://www.netb2b.com/

  - Advertising World
    http://advertising.utexas.edu/world/

  - NUA Internet Surveys
    http://www.nua.ie/surveys/

  - CNET builder.com
    http://www.builder.com

  - Webmonkey
    http://www.webmonkey.com

7.0 Site Search

Many commercial search engines offer a free "local search" downloads or shareware programs such as Harvest or ht://Dig that will act as a local search tool. One resource for finding these programs is Search Engine Watch at http://searchenginewatch.com/resources/software.html

You can also use a search solution that maintains and updates your site index externally. An example of this kind of tool is found at the Thunderstone site http://www.thunderstone.com/

- **Add Links to NASA Search**
  The goal of your site is to accommodate the user and provide the broadest possible set of tools to aid in searching information. To expand search options, consider adding a link to the Agency-wide search engine (NASA Search) so that a visitor can access information beyond your immediate site. If appropriate, include a link to the NASA Image Exchange (NIX) image search engine.

- **Add a Search Link on Every Page**
  Users enter sites at various locations (at the top or on sub-level pages, for example); to accommodate users, provide a link to the site search on every page to help user easily navigate the site. Also provide a link from the search page to a help page or advanced search page providing instructions on the syntax and examples for a particular search engine.

- **Make Headings Descriptive**
  Descriptive headings make it easy to index site contents for maximum search results. See the Web Marketing (Chapter 6.0) for more information on how to design site contents for search engines.

- **Use a Search Engine That Searches Metadata**
  Selecting a search engine that also searches against metadata keywords can help direct users to the most appropriate materials. The use of metadata keywords within a site will also accomplish the same for many Internet-wide search engines and assist in promoting a web site.

  A search engine that uses metadata descriptions in the results summary will allow control of the words displayed and aid a user on which link to follow. The use of metadata descriptions within a site will also accomplish the same for many Internet-wide search engines and assist in promoting a web site.

- **Keep the Search Engine Index Up-To-Date**
  The search engine index should reflect the updates made within the site. Users are frustrated when a search returns 404 errors for pages that no longer exist. Updating the index on a regular basis means that newly added pages will be found.

Tips

- New pages should be added at least weekly.
- Pages that can no longer be reached by the spider should be expired from the index after a month.
- Use appropriate wording on pages so that the search will find items. For example, never use an acronym on a page without spelling it out at least once. Consider the user who may not be familiar with the acronyms used by projects and programs.
- Use metadata or other invisible text on subsidiary pages to identify the topic being discussed, if this information is not included on the page.

- **Carefully Select Search Features**
  Consider the following features when choosing a site search mechanism:
  - Results sorted by relevance
  - Natural language queries
  - Search operators for greater control
  - Secondary searches within the current results
  - Search similar documents feature
  - Metadata keywords search feature
  - Metadata descriptions feature

Consider that standards for classification tags may change and, therefore, a search engine should be capable of customization to map any META tag into any searchable attribute.

- **Establish Filter Rules for Spiders**
  When managing a spider that gathers pages from your intranet, construct filtering rules to make sure that the spider indexes as much as possible of the content you want and ignores as much as possible of the content you don't want. Here are some reasons for filtering certain content:
  - Too much content to index (CD-ROM servers)
  - No useful content (a chain of automatically-generated calendar pages stretching way into the past or future)
  - Specialized bulk content that would skew search results (repositories of source code, large databases of machine parts)
  - Specialized repositories with their own search interfaces (image databases).

Every intranet requires different filtering rules for spidering to be successful. (The commercial Internet search engines use conservative rules that miss a lot of the content you want to index.) Additionally, you need to maintain content on a regular basis by

checking logs to see whether the spider is indexing undesirable content and checking the index to ensure that it has found the desired content. The ideal spidering strategy for an Intranet is to start at a single page and follow links everywhere in the applicable DNS domain. *Attempts to divide this domain into disjointed, separately-spiderable webs will almost certainly fail.*

- **Augment Search with a META Search Engine**
  Consider augmenting your search engine with a META search engine; your intranet will likely contain a number of sites with a large amount of content discoverable through the sites' own search interfaces. Rather than index all of it yourself, use a META search interface that will forward a query to many different search interfaces and return the results in real time. A discussion of META search products can be found on Search Engine Watch.

- **Use Robots**
  Webmasters who want to limit robot access to sensitive or private data on webservers can use the "robots.txt" file. A webmaster can easily restrict what directories and files a robot can access (HTML or otherwise) just by adding simple control parameters into the "robots.txt" file.

References:

- Search Engine Watch
  http://www.searchenginewatch.com

- Thunderstone
  http://www.thunderstone.com

- NASA Search
  http://198.116.142.33/97is.vts

- NASA Image Exchange
  http://nix.nasa.gov

- Xavatoria Search which uses Perl
  http://www.xav.com/scripts/xavatoria/

- Harvest Search Engine
  http://tardis.ed.ac.uk/harvest/

- ht://Dig
  http://www.htdig.org/

8.0 Authoring Tools

Because software versions are regularly updated, and because NASA cannot endorse commercial products (even those purchased), this section will ***not*** *discuss the pros and cons of particular applications. Rather, this section covers general principles for evaluating and purchasing software for editing web documents.*

As a Center, Headquarters has provided baseline software packages (http://www.hq.nasa.gov/help/hqwsstd.htm), including HTML and graphics editors. Most Center information technology security groups discourage downloading untested software, even demos from software publishers' reliable sites. When testing a product, it is best to coordinate with the Center IT group to obtain demonstration and test software.

8.1 Hypertext Markup Language (HTML)

Hypertext Markup Language (HTML) is a markup language that uses embedded tags that are read by a web browser and converted into a web page. The language has evolved since its invention in 1994, incorporating as standard tags many that were developed for specific browsers.

Tip: Use an HTML editor that recognizes the latest version and one version older of HTML, as defined by the World Wide Web consortium. As of 2000, the versions are HTML 3.2 and HTML 4.0. HTML 4.0 has been released, but many browsers do not yet fully support it. This is likely to continue to be the case with each new release.

HTML editors have blossomed from a few small applications that allowed only hand coding to more complex software with a graphical interface (also known as "what you see is what you get" or WYSIWYG). The WYSIWYG can make page creation much simpler; however, within the graphical interface there may be programming that effectively limits some options. For example, inserting graphics into a web page via "drag and drop" may generate relative Hypertext links that will have to be corrected before posting the page to a web server.

Some programs will allow switching back and forth between the graphical interface and HTML code, which can help overcome some of these limitations.

Tip: If considering a WYSIWYG editor, choose one that still allows editing of the HTML source code. Preference should be given to those that do not try to "second guess" what was intended, but leaves the code as is.

Tip: Learn HTML. If a web page does not appear as intended, WYSIWYG editors are no substitute for knowing HTML code.

The following are resources (from “World Wide Web Home Page Guidelines and Best Practices," prepared by the World Wide Web Federal Consortium, revised November 1996):

- Putting Information onto the Web - a collection of documents on authoring hypertext
  http://www.w3.org/hypertext/WWW/Provider/Overview.html#author

- NSF/NCSA World Wide Web Federal Consortium Training Materials Page
  http://hoohoo.ncsa.uiuc.edu:7777/consortium_info/agency_policy_guidelines/agency_policy.html

- Style Guide for On-line Hypertext; Tim Berners-Lee, World Wide Web Consortium
  http://www.w3.org/Provider/Style/Overview.html

- NCSA HTML Style Guide - National Center for Supercomputing Applications
  http://www.ncsa.uiuc.edu/styleguide

- Introduction to HTML and URLs; Ian Graham, University of Toronto
  http://www.utoronto.ca/webdocs/HTMLdocs/NewHTML/intro.html

- Guide to Writing HTML Documents; Dan LaLiberte, National Center for Supercomputing Applications
  http://www.hypernews.org/HyperNews/get/www/html/guides.html

- Composing Good HTML; Eric Tilton
  http://www.cs.cmu.edu/~tilt/cgh/

- Yale C/AIM WWW Style Manual; Patrick J. Lynch, Yale Center for Advanced Instructional Media
  http://info.med.yale.edu/caim/manual/contents.html

- Yahoo collection of WWW-related sites and documents
  http://www.yahoo.com/Computers_and_Internet/Internet/World_Wide_Web/

- Names and Addresses, URIs, URLs, URNs, URCs
  http://www.w3.org/Addressing/

8.2 Graphics

A variety of formats are available for graphics that can be incorporated into a web page. The two most commonly used are JPEG and GIF. Each allows for compression of images, using different techniques.

The important point in editing graphics for the web is to choose hardware and software that give the most flexibility on image resolution. Resolution is usually expressed as pixels per inch (ppi) or dots per inch (dpi). The higher the resolution, the larger the graphics file and the longer it takes to download.

A standard computer screen displays images at 72 dpi; higher resolutions are unnecessary for graphics meant to appear only on a monitor. Laser printers generally require higher resolution (120 dpi and above) and files meant for use in professional printing begin at 300 dpi and rise quickly. Any worthwhile graphics program will have the capability to change an image resolution, but it's best to start with a high-resolution graphic and reduce the resolution. Going from high to low resolution generally gives better results than going from low to high.

If high-resolution graphics need to be part of web site, these files should be clearly marked as large files. Small versions, called "thumbnails" should be posted to allow users to browse and select without having to download all the files. (See "Page Design" and "Site Design.")

Use graphics hardware and software that at minimum will create graphics in JPEG and GIF formats, which are viewable in both major commercial Web browsers and across platforms.

Note that web browsers may display fewer colors than are available in graphics applications. A graphic created with a 256-color graphics palette may appear slightly different in a web browser with a 216-color palette.

Use software that includes a web color palette, which displays graphics as they will appear on the web.

8.3 Plain Text Files

Plain text files appear in web browser windows exactly as described: plain, with no style, font or size variations. They can be useful for making large amounts of data accessible, but are more useful when transferred via the File Transfer Protocol (FTP) than when browsed on the web. In some cases, such as NASA HQ Press Releases, plain-text is a legacy from the early days of the web, when text files were easier to create than HTML and webmasters still had concerns about the download times of even small HTML files.

Limit the amount of plain text files on a web site, using HTML to take advantage of the Web's graphic capabilities.

Any word processing program should be capable of saving documents in a "text only" format with carriage returns (distinct from paragraph markers) at the end of each line.

Before converting a formatted document using a save function, convert the font to a fixed-width font (e.g., Courier) and adjust margins and spacing manually to get a better idea of how the document will look. This will make for fewer adjustments after the conversion. Bear in mind that most web users will choose their own fonts and sizes, so a document may appear differently on different browsers and user screens.

Tip: If a graphic is used to convey a concept and must not be altered, use a format like Adobe's Portable Document Format (pdf) to maintain the original look. These programs are also useful for print documents that must maintain the original formatting. You can also create a plain text file with a commonly available fixed-width font and margins that fit within the window of a standard 13-inch monitor. Maximum character width should be about 60 characters. Include a note at the top of the file indicating the font, size and margin width used in original formatting.

Tip: Always double-check documents within a standard web browser before posting. Some programs embed formatting commands that are not always deleted in conversions to text (e.g., "smart quotes" in Microsoft Word). These formatting commands then appear as indecipherable characters (Õ) in the browser.

8.4 Multimedia

The types of multimedia programs (sound, video, animation) available are too numerous to discuss in great detail. When incorporating multimedia into a site, bear in mind considerations of download time, audience interest, availability across platforms and accessibility requirements.

Use multimedia software that creates files viewable in the two major commercial web browsers without requiring an Internet user to purchase additional software. (Established software that requires users to download free "plug-ins" or "players" from reliable sites is acceptable as long as these plug-ins are available across platforms.)

Appendix A: Acronym List

Acronym	Description
CGI	Common Gateway Interface
CIO	Chief Information Officer
CSS	Cascading Style Sheets
DPI	Dots per Inch
DNS	Domain Name System
EFOIA	Electronic FOIA
FAQs	Frequently Asked Questions
FOIA	Freedom of Information Act
FTP	File Transfer Protocol
GIF	Graphics Interchange Format
HQ	Headquarters
HTML	HyperText Markup Language
https	HyperText Transmission Protocol, Secure
IP	Internet Protocol
IT	Information Technology
JPEG	Joint Photographic Experts Group
LDAP	Lightweight Directory Access Protocol
NASA	National Aeronautics and Space Administration
NASIRC	NASA Automated System Incident Response Capability
NHB	NASA Handbook
NIITA	NASA Integrated Information Technology Architecture
NMI	NASA Management Instruction
NPD	NASA Policy Directive
NPG	NASA Procedures and Guidelines
OMB	Office of Management and Budget
PC	Personal Computer
PDF	Portable Document Format
PKI	Public Key Infrastructure
POC	Point of Contact
QTVR	QuickTime Virtual Reality
SSI	Server Side Includes
SSL	Secure Socket Layer
URL	Universal Resource Locator
VRML	Virtual Reality Modeling Language
VPN	Virtual Private Network
WBP	Web Best Practices
WWW	World Wide Web
WYSIWYG	What You See Is What You Get
XML	eXtensible Markup Language

Appendix B: Use of the NASA Insignia and Seal on Agency Web Sites

Webmasters must use the NASA Seal, the NASA Insignia (known to employees as "the meatball") and the NASA Logotype (a.k.a., "the worm") as originally designed without alteration in shape, design, or color. Proper use of these "visual identifiers" is outlined in the Code of Federal Regulations (14 CFR §1221.1). The Code charges the NASA Graphics Coordinator in the Office of Public Affairs with ensuring the proper use of the NASA emblems in any setting in which they appear. At the time this document was written, the NASA Graphics Coordinator was Mr. Bert Ulrich, Public Services Division, Office of Public Affairs (phone: 202/358-1750).

The NASA Seal is reserved for use in association with the NASA Administrator. The NASA logotype has been retired, and should only be used in historical context or with the approval of the NASA Graphics Coordinator.

The Public Services Division has developed standards for using the NASA Insignia on the Web. A web site containing the guidelines and templates should be available in Fall 1998. A limited number of CD-ROMs containing templates and usage guidelines are available from the NASA Graphics Coordinator.
http://www.hq.nasa.gov/office/pao/insignia/text/Welcome.html

A few guidelines apply to use of the NASA Insignia:

- These emblems must not be used to imply the Agency's endorsement of another party's goods or services.

- They should not appear on non-NASA web pages. And NASA employees should not give permission for other groups to use the Insignia. This prohibition extends to contractors working directly on NASA projects and to organizational "partners" in joint projects. While some groups have appropriated the Insignia for their web pages, these uses have not been approved by the NASA Graphics Coordinator or the Office of the General Counsel. To resolve questions, contact Mr. Ulrich at 202/358-1750.

- They should never be used on NASA web sites as a "hot link" except to the NASA Home Page: http://www.nasa.gov/.

Appendix C: NASA Privacy Statement

This policy establishes how NASA will use information we gather about you from your visit to our web site. The privacy of our customers is of utmost importance to NASA.

If you visit a NASA site...

- **To read or download information:**

  We may collect and store information for statistical purposes. For example, we may count the number of visitors to the different sections of our site to help us make them more useful to visitors. Similar information is gathered for anonymous ftp, remote account login, or for other comparable types of connections.

- **To send us an E-mail:**

  By sending us an electronic mail message, you may be sending us personal information (e.g., name, address, E-mail address), as in an official Freedom of Information Act request. We may store the name and address of the requester in order to respond to the request or to otherwise resolve the subject matter of your E-mail.

- **To register:**

  Some of our sites ask visitors who request specific information to fill out a registration form. For example, vendors looking for marketing opportunities by visiting our Electronic Grants or NASA Acquisition Internet Service sites may be asked to “register” to obtain copies of Requests for Proposals or other NASA opportunities. Other information which may be collected at these sites through questionnaires, feedback forms, or other means, enable us to determine a visitor’s interests, with the goal of providing better service to our customers.

We want to be very clear: regardless of the information being transmitted to NASA, we will protect all such information consistent with applicable law.